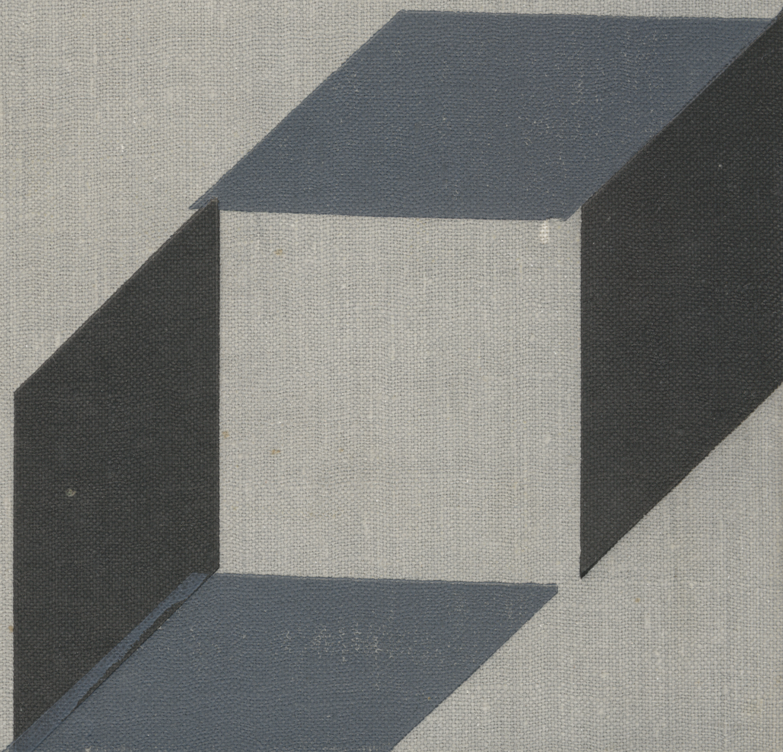
У. Прэтт

Цифровая обработка изображений

1



ИЗДАТЕЛЬСТВО
МИР

DIGITAL IMAGE PROCESSING

William K. Pratt

Professor of Electrical Engineering
Image Processing Institute
University of Southern California
Los Angeles, California

A WILEY–INTERSCIENCE PUBLICATION
JOHN WILEY AND SONS
NEW YORK/CHICHESTER/BRISBANE/TORONTO
1978

У. Прэтт

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

В 2-х книгах

1

Перевод с английского
под редакцией
канд. техн. наук Д. С. ЛЕБЕДЕВА

ИЗДАТЕЛЬСТВО «МИР»
МОСКВА 1982

ББК 32.94
П85
УДК 621.372.542

Прэтт У.

П85 Цифровая обработка изображений: Пер. с англ. — М.: Мир, 1982. — Кн.1 — 312 с., ил.

Монография крупного американского специалиста в области обработки изображений, который с 1971 по 1980 г. возглавлял Институт обработки изображений при Южнокалифорнийском университете в Лос-Анджелесе; наиболее полная по данной области знания.

В русском переводе монография выпускается в двух книгах. Книга 1 посвящена вопросам математического представления непрерывных и дискретных изображений. Подробно рассмотрены двумерные преобразования, в том числе преобразования Фурье, Адамара и Карунена–Лоэва.

Предназначена для специалистов, имеющих дело с обработкой изображений, а также для студентов, изучающих методы цифровой обработки изображений.

П $\frac{30401\text{-}186}{041(01)\text{-}82}$ 148-81, ч. 1 2405000000

ББК 32.94
П85

Редакция литературы по новой технике

ПРЕДИСЛОВИЕ К ПЕРЕВОДУ

Цифровая обработка изображений представляет собой самостоятельную область знания, которая быстро развивается во многих странах, в том числе и в Советском Союзе. Отечественной промышленностью выпускаются устройства ввода и регистрации изображений, дисплеи, графопостроители, специализированные процессоры и другое оборудование. Созданы мощные центры цифровой обработки изображений для исследования природных ресурсов и охраны окружающей среды, работают многочисленные цифровые установки для медицинской диагностики, в том числе томографические.

С обработкой изображений приходится иметь дело специалистам различного профиля, причем большинство из них закомятся с этим предметом самостоятельно. В связи с этим возникает необходимость в соответствующих учебных пособиях. По цифровой обработке изображений и смежным вопросам опубликовано много работ, но, как правило, они посвящены отдельным аспектам обработки изображений, что, естественно, затрудняет выбор литературы.

Монография Прэтта, несомненно, окажется полезной для советских специалистов, как знакомых с методами обработки изображений, так и для новичков. Она выгодно отличается от других книг на эту тему полнотой охвата материала. В ней рассматривается большинство известных методов и алгоритмов обработки изображений. Однако цифровая обработка изображений — предмет настолько обширный и разнородный, что уместить его в одну книгу совершенно невозможно. Так, в монографии не нашли отражения некоторые вопросы, отмеченные в предисловии автора, недостаточно полно рассмотрены статистические методы обработки изображений, которые в последние годы успешно применяются для улучшения и анализа изображений. Мы сочли целесообразным привести список статей и книг, опубликованных после выхода данной монографии в свет. В этот же список включены основные работы советских авторов, большой вклад которых в обработку изображений отражен в монографии совершенно недостаточно.

При переводе возникли большие терминологические трудности. Хотя не все решения, принятые редактором перевода и переводчиками, можно считать бесспорными, проделанная ими работа будет несомненно способствовать развитию и упорядочению терминологии в области цифровой обработки изображений.

Перевод выполнен В. П. Андреевым (гл. 17—20), А. Л. Зайцевым (гл. 4, 6, 7, 9, 10, 11), канд. техн. наук Д. С. Лебедевым (гл. 1—3), канд. техн. наук В. Г. Поляковым (гл. 21—24), Н. Н. Тетекиным (гл. 5, 8) и канд. техн. наук В. А. Хлебородовым (гл. 12—16).

Чл.-корр. АН СССР
В. И. Сифоров

ПРЕДИСЛОВИЕ

За последнее десятилетие область применения цифровой обработки изображений значительно расширилась. Этому способствовало повышение скорости работы, а также уменьшение стоимости и размеров цифровых вычислительных машин и технических средств обработки сигналов. Методы обработки изображений уже играют значительную роль в научных исследованиях, промышленности, медицине, космических исследованиях и информационных системах. Примерами применения этих методов могут служить цифровая передача изображений с космических кораблей, видеотелефонная связь по телефонным каналам, повышение четкости изображений, создаваемых электронным микроскопом, коррекция искажений изображений, принимаемых из космоса, автоматический анализ характера местности, исследование природных ресурсов по фотоснимкам, передаваемым со спутников Земли, формирование и улучшение качества биологических и медицинских изображений, включая рентгенограммы, термограммы и изображения радиоизотопной диагностики, автоматическое составление карт по аэрофотоснимкам, обнаружение дефектов в деталях машин с помощью промышленных рентгенограмм. Нет сомнения, что со временем методы обработки изображений найдут еще более широкое применение в медицине, во многих случаях облегчая врачу постановку диагноза, и в технике. Действия робота, наделенного «зрением», будут основываться на автоматическом анализе сцен. Эффективные методы кодирования изображений, по-видимому, позволят в будущем создать индивидуальные каналы двусторонней телевизионной связи как личного, так и служебного пользования.

Эта книга предназначается в качестве руководства по курсу «Цифровая обработка изображений» для студентов и аспирантов, специализирующихся в области электронной и вычислительной техники, и может быть также использована в качестве справочного пособия научными работниками и инженерами, связанными с исследованиями и практическими разработками в области обработки изображений. Цифровая обработка изображений — предмет широкого профиля, охватывающий вопросы физики, физиологии, электронной и вычислительной техники, математики. Предпола-

гается, что читатель имеет подготовку в какой-либо из этих областей в объеме программы технического вуза. Весьма желательно (но не обязательно) знакомство с теорией линейных систем, векторной алгеброй, теорией вероятностей и теорией случайных процессов.

Книга разделена на шесть частей. Часть 1, состоящая из трех глав, посвящена непрерывным изображениям. Здесь дано математическое представление непрерывного изображения, описаны психофизические свойства зрения человека и рассмотрены вопросы фотометрии и колориметрии. В части 2 приведены методы дискретизации и квантования, вопросы математического описания дискретных изображений и выбора критериев для оценки качества изображения. В части 3 обсуждаются методы двумерной обработки сигналов. Рассматриваются обобщенные линейные операторы, операторы псевдообращения, суперпозиции и свертки, а также операторы унитарных преобразований, в частности преобразований Фурье, Адамара и Карунена—Лоэва. В заключительной главе части 3 проводится сравнительный анализ линейных методов фильтрации, основанных на непосредственном вычислении свертки, выполнении преобразования Фурье и рекурсивной фильтрации.

В части 4 описываются методы улучшения и реставрации изображений. Для улучшения изображений используются операции, в результате которых изображение становится субъективно более приемлемым для зрительного восприятия либо облегчается визуальный или машинный анализ изображения. Задача реставрации состоит в том, чтобы повысить верность изображения, скорректировав искажения. В части 5 основное внимание сосредоточено на вопросах анализа сцен и интерпретации изображений. Рассматриваются задачи выделения и измерения признаков, обнаружения объектов, совмещения изображений, символического описания изображений, а также системы понимания изображений. Часть 6 посвящена вопросам кодирования изображений, осуществляемого с целью сокращения количества двоичных единиц, затрачиваемых на описание одноцветных и цветных изображений в системах передачи и запоминания изображений.

Хотя монография охватывает чрезвычайно широкий круг вопросов, многие темы, относящиеся к области обработки изображений, пришлось опустить из-за ограниченного объема книги. Среди них наиболее важными являются распознавание и классификация образов, цифровая голография и томография. Библиография по этим вопросам дана в приложении 1.

У. К. Прэтт

Малибу, Калифорния
январь, 1978

БЛАГОДАРНОСТИ

Исследования по вопросам обработки изображений начались в Южнокалифорнийском университете (ЮКУ) в 1962 г. с очень скромных по своим масштабам работ. Однако с ростом интереса к этой области во всем мире они приобрели широкий размах. В 1971 г. д-р З. Капрелян, декан Инженерно-технической школы и вице-президент университета по вопросам научных исследований и административного управления, объявил о создании при университете Института обработки изображений. В настоящее время этот институт — один из крупнеших в мире исследовательских центров по данному кругу проблем, оснащенный первоклассным оборудованием. Все эти обстоятельства имеют непосредственное отношение к появлению данной монографии. Я признателен д-ру Капреляну за его роль в организации поддержки университетом исследований в области обработки изображений. Хочется поблагодарить членов ученого совета института, оказавших мне неоценимую помощь при подготовке рукописи, а именно Г. Эндрюса, Л. Дэвиссона, В. Фрэя, А. Хабиби, Э. Холла, Р. Гершеля, А. Джейна, Р. Крюгера, Н. Нахи, Р. Невейшиа, Г. Робинсон, А. Савчука, В. Томпсона и Л. Уэлча. Я благодарен также моим студентам И. Абду, Б. Ашьяри, В. Чену, Ф. Даваряну, М. Юнсу, К. Мэнсиллу, Н. Мацаренасу, Дж. Роузу, К. Ридеру и Р. Уэллису за техническую помощь при подготовке книги.

Монография написана на основе конспекта лекций по курсу «Обработка изображений», читавшихся на факультете электротехники ЮКУ. Я признателен многим студентам за их замечания, сделанные по первым вариантам этого курса. Хочется отметить также работу по улучшению конспекта, проделанную профессорами ЮКУ Г. Эндрюсом, Э. Холлом, В. Фрэем и А. Хабиби, которые использовали этот конспект в процессе преподавания курса обработки изображений.

Подготовка любой книги чрезвычайно трудная задача. Мне очень повезло, что рукопись перепечатали сотрудницы ЮКУ Флоренс Теббетс, Джойс Сеге и Мэрилин Чен. Р. Шмидт и его коллеги из Института обработки изображений выполнили необходимые фотографические работы; что касается графических

материалов, то заслуга их подготовки принадлежит Д. Хауленду из ЮКУ. Приведенные в книге примеры изображений, полученных с помощью ЭВМ, — это результат большой работы С. Джонсона и Э. Казаняна по составлению вычислительных программ. Фотографическая регистрация ряда изображений по цифровым данным, полученным на ЭВМ, выполнена д-ром К. Ридером (Ford Aerospace and Communications Corp., Пало-Альто, Калифорния). Всем этим лицам я выражаю свою искреннюю благодарность.

Наконец, я очень признателен Агентству перспективных исследовательских работ (ARPA) и Бюро методов обработки информации, которое входит в его состав, за финансовую помощь, а также персоналу Лаборатории измерений ВВС (контракт №F-08606-72-C-0008) и военно-воздушной базы Райт-Паттерсон (контракт №F-33615-76-C-1203) за проведение контрольных измерений.

У. К. П.

Часть 1

Непрерывные изображения

В этой книге рассматриваются главным образом цифровые, а не аналоговые методы обработки изображений. Однако в большинстве случаев цифровое изображение получают из естественного непрерывного изображения. Исключением являются искусственные изображения, например созданные с помощью ЭВМ испытательные таблицы или томограммы. Следовательно, важно понимать, как происходит формирование изображения с помощью различных преобразователей и оптических систем, включая зрительную систему человека. Другой существенный вопрос — световые измерения для количественного описания изображений. Полезно также изучить пространственные и временные свойства непрерывного светового поля, которые обусловливают характер взаимосвязей между элементами цифрового изображения. Эти вопросы рассматриваются в гл. 1—3. Глава 1 посвящена математическому описанию непрерывных изображений, в гл. 2 описываются психофизические свойства зрения, а в гл. 3 приводятся сведения по фотометрии и колориметрии.

Глава 1

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ

При разработке и анализе систем обработки изображений удобно, а часто и необходимо иметь математическое описание подлежащих обработке изображений. Существуют два основных подхода к такому описанию: детерминированный и статистический. При детерминированном описании вводится математическая функция, представляющая изображение, и рассматриваются свойства изображения в каждой его точке. При статистическом описании изображение определяется усредненными характеристиками. Следующие разделы посвящены как детерминированному, так и статистическому описанию непрерывных изображений. Хотя приведенный анализ относится к визуально воспринимаемым изображениям, многие результаты могут быть распространены на двумерные изменяющиеся во времени сигналы и поля общего вида.

1.1. ПРЕДСТАВЛЕНИЕ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ

Распределение энергии источника светового излучения по пространственным координатам x, y, времени t и длинам волн λ описывается функцией $C(x, y, t, \lambda)$. Энергия излучения пропорциональна квадрату амплитуды электрического поля и поэтому представляет собой действительную положительную величину. В изображающих системах максимальная яркость изображения ограничена из-за насыщения светочувствительной пленки или перегревания люминофора кинескопа. Следовательно,

$$0 \leqslant C(x, y, t, \lambda) \leqslant A, \tag{1.1.1}$$

где A — максимальная яркость изображения. Размеры изображения ограничены формирующей системой и средой, на которую оно записывается. В целях упрощения будем считать, что все изображения отличны от нуля только в прямоугольной области, для которой

$$-L_x \leqslant x \leqslant L_x, \quad (1.1.2а)$$

$$-L_y \leqslant y \leqslant L_y. \quad (1.1.2б$$

Изображение наблюдается в течение конечного промежутка времени, т. е.

$$-T \leqslant t \leqslant T. \tag{1.1.2в}$$

Таким образом, величина $C(x, y, t, \lambda)$ является ограниченной функцией четырех ограниченных переменных. Будем считать эту функцию непрерывной в области ее определения.

Ощущение светлоты, возникающее в зрительной системе человека, обычно определяется мгновенной яркостью светового поля, т. е. величиной

$$Y(x, y, t) = \int_0^\infty C(x, y, t, \lambda)\, V_S(\lambda)\, d\lambda, \tag{1.1.3}$$

где $V_S(\lambda)$ — спектральная чувствительность человеческого зрения. Цветовые ощущения можно описать набором так называемых координат цвета, пропорциональных интенсивностям красного, зеленого и синего цвета, смесь которых дает заданный цвет. Для произвольной красно-зелено-синей координатной системы текущие значения координат цвета равны

$$R(x, y, t) = \int_0^\infty C(x, y, t, \lambda)\, R_S(\lambda)\, d\lambda, \tag{1.1.4а}$$

$$G(x, y, t) = \int_0^\infty C(x, y, t, \lambda)\, G_S(\lambda)\, d\lambda, \tag{1.1.4б}$$

$$B(x, y, t) = \int_0^\infty C(x, y, t, \lambda)\, B_S(\lambda)\, d\lambda, \tag{1.1.4в}$$

где $R_S(\lambda)$, $G_S(\lambda)$, $B_S(\lambda)$ — удельные координаты для набора основных цветов — красного, зеленого и синего, равные координатам цвета монохроматического излучения единичной интенсивности с длиной волны λ. В спектрозональных системах изображение для i-й спектральной зоны описывается выражением

$$F_i(x, y, t) = \int_0^\infty C(x, y, t, \lambda)\, S_i(\lambda)\, d\lambda, \tag{1.1.5}$$

где $S_i(\lambda)$ — спектральная чувствительность i-го датчика.

Для простоты будем во всех случаях описывать изображение, сформированное некоторой физической системой, с помощью функции $F(x, y, t)$. Для одноцветной системы функция $F(x, y, t)$ представляет распределение яркости или какой-либо другой физической величины, связанной с яркостью. Для системы воспроизведения цветных изображений $F(x, y, t)$ есть одна из координат цвета. Функция $F(x, y, t)$ используется также для описа-

ния других полей, например меняющегося во времени пространственного распределения шума видеодатчика.

В соответствии с обычным определеним среднего значения одномерного сигнала средняя по времени яркость изображения в данной точке x, y определяется как

$$\langle F(x, y, t)\rangle_T = \lim_{T\to\infty}\left\{(1/2T)\int_{-T}^{T} F(x, y, t)\, L(t)\, dt\right\}, \qquad (1.1.6)$$

где $L(t)$ — временная весовая функция. Аналогично определяется пространственная средняя яркость в момент времени t:

$$\langle F(x, y, t)\rangle_S = \lim_{\substack{L_x\to\infty \\ L_y\to\infty}}\left\{(1/4L_xL_y)\int_{-L_x}^{L_x}\int_{-L_y}^{L_y} F(x, y, t)\, dx\, dy\right\}. \qquad (1.1.7)$$

Во многих системах воспроизведения изображений, например в проекционных устройствах, изображение не изменяется во времени и переменная t может быть опущена. В системах другого типа, например кинематографических, аргумент t меняется дискретно. В дальнейшем аргумент t функции, описывающей изображения, будет обычно опускаться.

1.2. ДВУМЕРНЫЕ СИСТЕМЫ

В двумерной системе набор исходных двумерных функций $F_1(x, y)$, $F_2(x, y)$, ..., $F_N(x, y)$ отображается в набор двумерных функций $G_1(x, y)$, $G_2(x, y)$, ..., $G_M(x, y)$, где x, y — независимые непрерывные пространственные переменные, аргументы этих функций ($\infty < x, y < \infty$). Это отображение может быть задано набором операторов $\mathcal{O}_m\{\cdot\}$ при $m = 1, 2, \ldots, M$, которые связывают входные и выходные функции:

$$\begin{gathered} G_1(x, y) = \mathcal{O}_1\{F_1(x, y), \ldots, F_N(x, y)\}, \\ \cdots\cdots\cdots\cdots\cdots \\ G_M(x, y) = \mathcal{O}_M\{F_1(x, y), \ldots, F_N(x, y)\}. \end{gathered} \qquad (1.2.1)$$

Число входных функций N может быть больше, меньше или равняться числу выходных. При $N = M = 1$

$$G(x, y) = \mathcal{O}\{F(x, y)\}. \qquad (1.2.2)$$

В одномерных физических системах, в которых независимой переменной является время, выходной сигнал представляет собой функцию прошлых и настоящих значений входного сигнала, но не может быть функцией будущих значений. Такие системы, называемые физически реализуемыми, удовлетворяют принципу причинности. Двумерные системы в общем случае не обладают

этим свойством: пространственные переменные x, y могут принимать отрицательные значения по отношению к некоторому началу координат.

Чтобы продолжить обсуждение свойств двумерных систем, необходимо познакомиться с некоторыми специальными видами операторов.

1.3. СИНГУЛЯРНЫЕ ОПЕРАТОРЫ

Сингулярные операторы широко применяются при анализе двумерных систем, особенно систем, в которых производится дискретизация непрерывных функций. Двумерная дельта-функция Дирака есть сингулярный оператор, обладающий следующими свойствами:

$$\delta(x, y) = \begin{cases} \infty, & x = 0,\ y = 0, \\ 0 & \text{в остальных случаях,} \end{cases} \tag{1.3.1а}$$

$$\delta(x - \xi, y - \eta) = \begin{cases} \infty, & x = \xi,\ y = \eta, \\ 0 & \text{в остальных случаях,} \end{cases} \tag{1.3.1б}$$

$$\int_{-\varepsilon}^{\varepsilon}\int \delta(x, y)\, dx\, dy = 1 \text{ при } \varepsilon > 0, \tag{1.3.1в}$$

$$\int_{-\infty}^{\infty}\int F(\xi, \eta)\, \delta(x - \xi, y - \eta)\, d\xi\, d\eta = F(x, y). \tag{1.3.1г}$$

Величина ε здесь обозначает бесконечно малый предел интегрирования.

Двумерная дельта-функция может быть представлена как произведение двух одномерных дельта-функций ортогональных координат x, y:

$$\delta(x, y) = \delta(x)\, \delta(y), \tag{1.3.2}$$

где одномерные дельта-функции удовлетворяют одномерным соотношениям, аналогичным (1.3.1). Дельта-функция может быть также определена как предел некоторых функций [1, стр. 275], например прямоугольной функции

$$\delta(x, y) = \lim_{\alpha\to\infty} [\alpha^2 \operatorname{rect}\{\alpha x\} \operatorname{rect}\{\alpha y\}], \tag{1.3.3а}$$

круговой функции

$$\delta(x, y) = \lim_{\alpha\to\infty} [(\alpha^2/\pi) \operatorname{cirk}\{\alpha \sqrt{x^2 + y^2}\}], \tag{1.3.3б}$$

гауссовой функции

$$\delta(x, y) = \lim_{\alpha\to\infty} [\alpha^2 \exp\{-\alpha^2\pi(x^2 + y^2)\}], \tag{1.3.3в}$$

sinc-функции

$$\delta(x, y) = \lim_{\alpha\to\infty} [\alpha^2 \operatorname{sinc}\{\alpha x\} \operatorname{sinc}\{\alpha y\}], \tag{1.3.3г}$$

бесселевой функции

$$\delta(x, y) = \lim_{\alpha\to\infty}\left[\frac{\alpha J_1(2\pi\alpha\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}\right], \tag{1.3.3д}$$

где

$$\operatorname{rect}\{x\} = \begin{cases} 1, & |x| \leqslant 1/2, \\ 0, & |x| > 1/2. \end{cases} \tag{1.3.4а}$$

$$\operatorname{circ}\{r\} = \begin{cases} 1, & r \leqslant 1, \\ 0, & r > 1, \end{cases} \tag{1.3.4б}$$

$$\operatorname{sinc}\{x\} = \sin \pi x/\pi x. \tag{1.3.4в}$$

Другим полезным определением дельта-функции является следующее тождество [2, стр. 99]:

$$\delta(x-\xi, y-\eta) = (1/4\pi^2)\iint_{-\infty}^{\infty} \exp\{i[u(x-\xi) + v(y-\eta)]\}\, du\, dv, \tag{1.3.5}$$

где $i = \sqrt{-1}$.

1.4. ЛИНЕЙНЫЕ ОПЕРАТОРЫ

Двумерная система называется линейной, если для нее справедлив принцип суперпозиции. В частном случае отображения функции в функцию для этого требуется, чтобы

$$\mathcal{O}\{a_1F_1(x, y) + a_2F_2(x, y)\} = a_1\mathcal{O}\{F_1(x, y)\} + a_2\mathcal{O}\{F_2(x, y)\}, \tag{1.4.1}$$

где a_1, a_2 — некоторые постоянные (могут быть комплексными). Определение свойства суперпозиции можно легко распространить на отображение (1.2.1) общего вида.

Используя свойство дельта-функции (1.3.1г), функцию на входе системы $F(x, y)$ можно представить как взвешенную сумму дельта-функций:

$$F(x, y) = \iint_{-\infty}^{\infty} F(\xi, \eta)\, \delta(x-\xi, y-\eta)\, d\xi\, d\eta, \tag{1.4.2}$$

где $F(\xi, \eta)$ — весовой множитель дельта-импульса, имеющего координаты ξ, η на плоскости (x, y) (рис. 1.4.1). Если функция на выходе линейной системы

$$G(x, y) = \mathcal{O}\{F(x, y)\}, \tag{1.4.3}$$

то

$$G(x, y) = \mathcal{O}\left\{\iint\limits_{-\infty}^{\infty} F(\xi, \eta)\,\delta(x-\xi, y-\eta)\,d\xi\,d\eta\right\}, \quad (1.4.4а)$$

или

$$G(x, y) = \iint\limits_{-\infty}^{\infty} F(\xi, \eta)\,\mathcal{O}\{\delta(x-\xi, y-\eta)\}\,d\xi\,d\eta. \quad (1.4.4б)$$

Для перехода от выражения (1.4.4а) к (1.4.4б) был изменен порядок выполнения операций линейного преобразования и интегрирования. Линейный оператор действовал только на тот множитель подынтегрального выражения (1.4.4а), который зависит от пространственных переменных x, y. Запишем второй множитель подынтегрального выражения (1.4.4б) как

$$H(x, y; \xi, \eta) \equiv \mathcal{O}\{\delta(x-\xi, y-\eta)\}. \quad (1.4.5)$$

Будем называть эту функцию импульсным откликом двумерной системы. Импульсный отклик оптической системы часто назы-

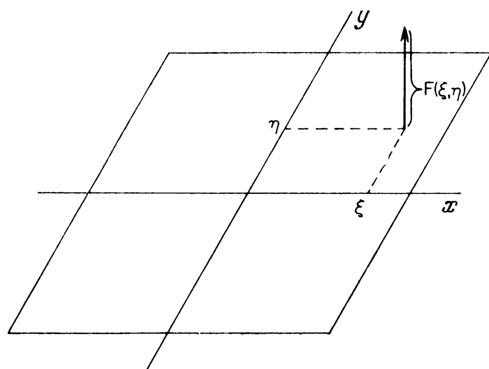


Рис. 1.4.1. Представление функции, описывающей изображение, в виде суперпозиции дельта-функций.

вается функцией рассеяния точки. Подстановка импульсного отклика в соотношение (1.4.4б) дает интеграл суперпозиции

$$G(x, y) = \iint\limits_{-\infty}^{\infty} F(\xi, \eta)\,H(x, y; \xi, \eta)\,d\xi\,d\eta. \quad (1.4.6)$$

Линейная двумерная система называется пространственно-инвариантной (изопланатической), если ее импульсный отклик зависит только от разностей координат $x - \xi$, $y - \eta$. Для оптической системы, показанной на рис. 1.4.2, это значит, что при перемещении точечного источника в предметной плоскости изображение этого

источника в плоскости фокусировки будет также изменять положение, но сохранять форму. Для пространственно-инвариантной системы

$$H(x, y; \xi, \eta) = H(x - \xi, y - \eta) \qquad (1.4.7)$$

и интеграл суперпозиции имеет особую форму, называемую интегралом свертки:

$$G(x, y) = \iint_{-\infty}^{\infty} F(\xi, \eta) H(x - \xi, y - \eta) d\xi d\eta. \qquad (1.4.8а)$$

Операция свертки символически записывается как

$$G(x, y) = F(x, y) * H(x, y). \qquad (1.4.8б)$$

Интеграл свертки симметричен, т. е.

$$G(x, y) = \iint_{-\infty}^{\infty} F(x - \xi, y - \eta) H(\xi, \eta) d\xi d\eta. \qquad (1.4.9)$$

Процесс свертки иллюстрируется на рис. 1.4.3. На рис. 1.4.3, *а* и 1.4.3, *б* изображены функция $F(x, y)$ на входе и импульсный

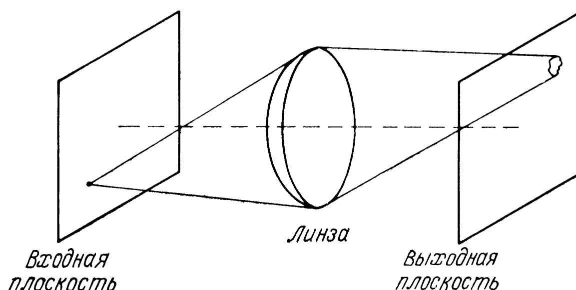


Рис. 1.4.2. Изображение точечного источника света в оптической системе.

отклик. На рис. 1.4.3, *в* показан импульсный отклик при обращении координат, а на рис. 1.4.3, *г* — со сдвигом на величину x, y. На рис. 1.4.3, *д* заштрихована область, в которой произведение $F(\xi, \eta) H(x - \xi, y - \eta)$, входящее в подынтегральное выражение (1.4.8, *а*), не равно нулю. Интегрирование по этой области дает величину $G(x, y)$ для заданных значений координат x, y. Таким образом, функция $G(x, y)$ на выходе может быть найдена сканированием входной функции скользящим «окном» — обращенным импульсным откликом, и интегрированием по области, в которой эти функции перекрываются.

1.5. ДИФФЕРЕНЦИАЛЬНЫЕ ОПЕРАТОРЫ

Обнаружение и выделение контуров (границ) на изображении, т. е. точек с резкими перепадами яркости, обычно осуществляются дифференцированием функции $F(x, y)$ по пространственным координатам и сравнением результата дифференцирования с порогом. Первые производные по пространственным координатам x и y определяются как

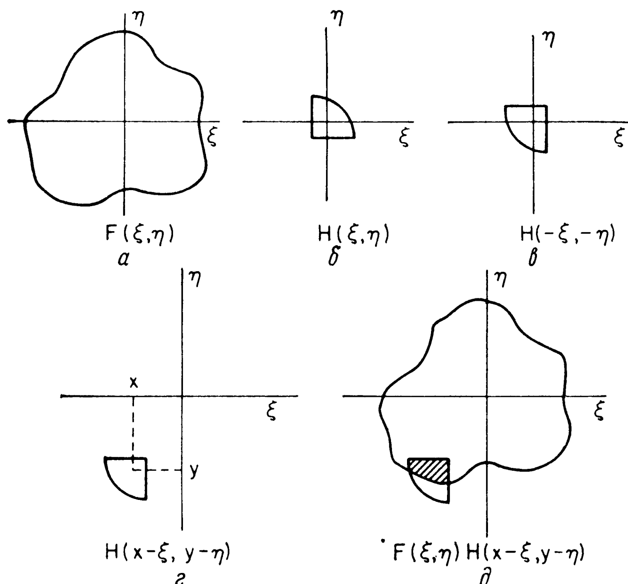


Рис. 1.4.3. Пример двумерной свертки.

$$d_x = \partial F(x, y)/\partial x, \quad (1.5.1а)$$

$$d_y = \partial F(x, y)/\partial y. \quad (1.5.1б)$$

Производная по направлению z, составляющему угол Φ с горизонтальной осью, по определению [3, стр. 10] равна

$$\nabla\{F(x, y)\} = \partial F(x, y)/\partial z = d_x \cos\Phi + d_y \sin\Phi. \quad (1.5.2)$$

Модуль градиента есть

$$|\nabla\{F(x, y)\}| = \sqrt{d_x^2 + d_y^2}. \quad (1.5.3)$$

Вторые производные по пространственным координатам x и y по определению равны

$$d_{xx} = \partial^2 F(x, y)/\partial x^2, \quad (1.5.4а)$$

$$d_{yy} = \partial^2 F(x, y)/\partial y^2. \quad (1.5.4б)$$

Сумма производных (1.5.4) дает лапласиан

$$L\{F(x, y)\} = d_{xx} + d_{yy} = \partial^2 F(x, y)/\partial x^2 + \partial^2 F(x, y)/\partial y^2. \quad (1.5.5)$$

Следует отметить, что лапласиан является скалярной величиной, которая не зависит от направления на плоскости x, y, тогда как градиент есть зависящая от направления векторная величина.

1.6. ДВУМЕРНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

В результате двумерного преобразования Фурье функции $F(x, y)$, описывающей изображение, получается спектр этого изображения, который определяется как

$$\mathcal{F}(\omega_x, \omega_y) = \iint_{-\infty}^{\infty} F(x, y) \exp\{-i(\omega_x x + \omega_y y)\}\, dx\, dy, \quad (1.6.1)$$

где ω_x, ω_y — пространственные частоты, а $i = \sqrt{-1}$. Если обозначить оператор преобразования Фурье через $\mathcal{O}_F\{\cdot\}$, то можно записать

$$\mathcal{F}(\omega_x, \omega_y) = \mathcal{O}_F\{\mathcal{F}(x, y)\}. \quad (1.6.2)$$

В общем случае спектр $\mathcal{F}(\omega_x, \omega_y)$ есть комплексная величина. Его можно разложить на действительную и мнимую части:

$$\mathcal{F}(\omega_x, \omega_y) = \mathcal{R}(\omega_x, \omega_y) + i\mathcal{I}(\omega_x, \omega_y) \quad (1.6.3а)$$

или представить с помощью амплитуды и фазы:

$$\mathcal{F}(\omega_x, \omega_y) = \mathcal{M}(\omega_x, \omega_y) \exp\{i\phi(\omega_x, \omega_y)\}, \quad (1.6.3б)$$

где

$$\mathcal{M}(\omega_x, \omega_y) = [\mathcal{R}^2(\omega_x, \omega_y) + \mathcal{I}^2(\omega_x, \omega_y)]^{1/2}, \quad (1.6.4а)$$

$$\varphi(\omega_x\ \omega_y) = \operatorname{arctg}\{\mathcal{I}(\omega_x, \omega_y)/\mathcal{R}(\omega_x, \omega_y)\}. \quad (1.6.4б)$$

Достаточным условием существования фурье-спектра функции $F(x, y)$ является абсолютная интегрируемость этой функции, т. е. условие

$$\iint_{-\infty}^{\infty} |F(x, y)|\, dx\, dy < \infty. \quad (1.6.5)$$

Исходная функция $F(x, y)$ может быть восстановлена обратным преобразованием Фурье:

$$F(x, y) = (1/4\pi^2) \iint_{-\infty}^{\infty} \mathcal{F}(\omega_x, \omega_y) \exp\{i(\omega_x x + \omega_y y)\}\, d\omega_x\, d\omega_y. \quad (1.6.6а)$$

Это соотношение в операторной форме можно записать как

$$F(x, y) = \mathcal{O}_{\mathcal{F}}^{-1}\{\mathcal{F}(\omega_x, \omega_y)\}. \quad (1.6.6б)$$

Поскольку ядро двумерного преобразования Фурье разделимо, это преобразование может быть выполнено в два этапа. Сначала находится

$$\mathcal{F}_y(\omega_x, y) = \int_{-\infty}^{\infty} F(x, y) \exp\{-i\omega_x x\}\, dx, \tag{1.6.7}$$

а затем

$$\mathcal{F}(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \mathcal{F}_y(\omega_x, y) \exp\{-i\omega_y y\}\, dy. \tag{1.6.8}$$

Ниже приводятся несколько полезных свойств двумерного преобразования Фурье. Их доказательства можно найти в книгах [1, 2].

Функциональные свойства

Если функция $F(x, y)$ разделима по пространственным переменным, так что

$$F(x, y) = f_x(x) f_y(y), \tag{1.6.9}$$

то

$$\mathcal{F}(\omega_x, \omega_y) = \mathscr{f}_x(\omega_x) \mathscr{f}_y(\omega_y), \tag{1.6.10}$$

где $\mathscr{f}_x(\omega_x)$ $\mathscr{f}_y(\omega_y)$ — одномерные фурье-спектры функций $f_x(x)$, $f_y(y)$. Если $\mathcal{F}(\omega_x, \omega_y)$ есть фурье-спектр функции $F(x, y)$, то $\mathcal{F}^*(-\omega_x, -\omega_y)$ является фурье-спектром функции $F^*(x, y)$. (Звездочка обозначает комплексную сопряженность.) Если функция $F(x, y)$ симметрична, т. е. $F(x, y) = F(-x, -y)$, то $\mathcal{F}(\omega_x, \omega_y) = \mathcal{F}(-\omega_x, -\omega_y)$.

Линейность

Оператор преобразования Фурье линеен:

$$\mathcal{O}_{\mathcal{F}}\{aF_1(x, y) + bF_2(x, y)\} = a\mathcal{F}_1(\omega_x, \omega_y) + b\mathcal{F}_2(\omega_x, \omega_y), \tag{1.6.11}$$

где a, b — постоянные.

Изменение масштаба

Изменение масштаба пространственных переменных приводит к обратному изменению масштаба пространственных частот и пропорциональному изменению значений спектра:

$$\mathcal{O}_{\mathcal{F}}\{F(ax, by)\} = (1/|ab|)\, \mathcal{F}(\omega_x/a, \omega_y/b). \tag{1.6.12}$$

Следовательно, сжатие вдоль одной из осей плоскости (x, y) приводит к растяжению вдоль соответствующей оси частотной плоскости и наоборот. Происходит также пропорциональное изменение значений спектра.

Сдвиг

Сдвиг (изменение координат) на исходной плоскости приводит к фазовым изменениям на частотной плоскости:

$$\mathcal{O}_{\mathcal{F}}\{F(x-a,\ y-b)\} = \mathcal{F}(\omega_x,\ \omega_y)\exp\{-i(\omega_x a + \omega_y b)\}. \quad (1.6.13а)$$

Наоборот, сдвиг на частотной плоскости вызывает фазовые изменения исходной функции:

$$\mathcal{O}_{\mathcal{F}}^{-1}\{\mathcal{F}(\omega_x - a_x,\ \omega_y - b_y)\} = F(x,\ y)\exp\{i(a_x x + b_y y)\}. \quad (1.6.13б)$$

Свертка

Фурье-спектр функции, полученный в результате свертки двух функций, равен произведению спектров исходных функций:

$$\mathcal{O}_{\mathcal{F}}\{F(x,\ y) * H(x,\ y)\} = \mathcal{F}(\omega_x,\ \omega_y)\,\mathcal{H}(\omega_x,\ \omega_y). \quad (1.6.14)$$

Обратная теорема утверждает, что

$$\mathcal{O}_{\mathcal{F}}\{F(x,\ y)\,H(x,\ y)\} = (1/4\pi^2)\,\mathcal{F}(\omega_x,\ \omega_y) * \mathcal{H}(\omega_x,\ \omega_y). \quad (1.6.15)$$

Теорема Парсеваля

Два представления энергии изображения — через функцию $F(x,\ y)$ и фурье-спектр $\mathcal{F}(\omega_x,\ \omega_y)$ — связаны следующим образом:

$$\iint_{-\infty}^{\infty} |F(x,\ y)|^2\,dx\,dy = (1/4\pi^2)\iint_{-\infty}^{\infty} |\mathcal{F}(\omega_x,\ \omega_y)|^2\,d\omega_x\,d\omega_y. \quad (1.6.16)$$

Теорема о спектре автокорреляционной функции

Фурье-спектр двумерной автокорреляционной функции изображения равен квадрату модуля фурье-спектра этого изображения:

$$\mathcal{O}_{\mathcal{F}}\left\{\iint_{-\infty}^{\infty} F(\alpha,\ \beta)\,F^*(\alpha - x,\ \beta - y)\,d\alpha\,d\beta\right\} = |\mathcal{F}(\omega_x,\ \omega_y)|^2. \quad (1.6.17)$$

Спектры пространственных производных

Фурье-спектры первых пространственных производных функции $F(x,\ y)$ связаны с ее фурье-спектром следующими соотношениями:

$$\mathcal{O}_{\mathcal{F}}\{\partial F(x,\ y)/\partial x\} = -i\,\omega_x\mathcal{F}(\omega_x,\ \omega_y), \quad (1.6.18а)$$

$$\mathcal{O}_{\mathcal{F}}\{\partial F(x,\ y)/\partial y)\} = -i\omega_y\mathcal{F}(\omega_x,\ \omega_y). \quad (1.6.18б)$$

Следовательно, спектр лапласиана равен

$$\mathcal{O}_{\mathcal{F}}\{\partial^2 F(x,y)/\partial x^2 + \partial^2 F(x,\ y)/\partial y^2\} = -(\omega_x^2 + \omega_y^2)\,\mathcal{F}(\omega_x,\ \omega_y). \quad (1.6.19)$$

1.7. АНАЛИЗ ЛИНЕЙНЫХ СИСТЕМ С ПОМОЩЬЮ ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Теорема о преобразовании Фурье свертки (1.6.14) оказывается очень полезным средством при анализе линейных систем. Рассмотрим функцию $F(x, y)$, описывающую изображение на входе линейной системы с импульсным откликом $H(x, y)$. Изображение на выходе описывается функцией $G(x, y)$, получаемой в результате свертки:

$$G(x, y) = \iint\limits_{-\infty}^{\infty} F(\alpha, \beta) H(x-\alpha, y-\beta)\, d\alpha\, d\beta. \tag{1.7.1}$$

Выполнив преобразование Фурье обеих частей этого равенства и поменяв порядок интегрирования в его правой части, получим

$$\mathcal{G}(\omega_x, \omega_y) = \iint\limits_{-\infty}^{\infty} F(\alpha, \beta) \left[\iint\limits_{-\infty}^{\infty} H(x-\alpha, y-\beta) \times \exp\{-i(\omega_x x + \omega_y y)\}\, dx\, dy \right] d\alpha\, d\beta. \tag{1.7.2}$$

Согласно теореме о сдвиге (1.6.13), внутренний интеграл равен произведению спектра функции $H(x, y)$ и экспоненциального множителя фазового сдвига. Поэтому

$$\mathcal{G}(\omega_x, \omega_y) = \iint\limits_{-\infty}^{\infty} F(\alpha, \beta)\, \mathcal{H}(\omega_x, \omega_y) \exp\{-i(\omega_x \alpha + \omega_y \beta)\}\, d\alpha\, d\beta. \tag{1.7.3}$$

Выполнив преобразование Фурье, получим

$$\mathcal{G}(\omega_x, \omega_y) = \mathcal{H}(\omega_x, \omega_y)\, \mathcal{F}(\omega_x, \omega_y). \tag{1.7.4}$$

Наконец, обратное преобразование Фурье дает функцию, описывающую изображение на выходе:

$$G(x, y) = (1/4\pi^2) \iint\limits_{-\infty}^{\infty} \mathcal{H}(\omega_x, \omega_y)\, \mathcal{F}(\omega_x, \omega_y) \exp\{i(\omega_x x + \omega_y y)\}\, d\omega_x\, d\omega_y. \tag{1.7.5}$$

Выражения (1.7.1) и (1.7.5) представляют два альтернативных способа определения выходного изображения линейной пространственно-инвариантной системы. Выбор того или иного подхода зависит обычно от решаемой задачи.

1.8. ОБОБЩЕННЫЕ ЛИНЕЙНЫЕ СИСТЕМЫ

В разд. 1.4 были введены понятия линейности и суперпозиции. Оппенгейм [4] ввел обобщенную суперпозицию с тем, чтобы распространить понятие линейности на более широкий класс систем.

Рассмотрим две функции, описывающие изображения, $F_1(x, y)$ и $F_2(x, y)$, которые, взаимодействуя некоторым образом ($\Diamond$), дают функцию $G(x, y)$:

$$G(x, y) = F_1(x, y) \Diamond F_2(x, y). \tag{1.8.1}$$

Пусть $\mathcal{O}_G\{\cdot\}$ — оператор системы, преобразующей $G(x, y)$, который обладает следующими свойствами:

$$\mathcal{O}_G\{F_1(x, y) \Diamond F_2(x, y)\} = \mathcal{O}_G\{F_1(x, y)\} \Diamond \mathcal{O}_G\{F_2(x, y)\} \tag{1.8.2а}$$

и

$$\mathcal{O}_G\{k : F(x, y)\} = k : \mathcal{O}_G\{F(x, y)\}, \tag{1.8.2б}$$

где k — постоянная, а двоеточие обозначает обобщенное умножение на постоянную. В работе [4] показано, что если операция $\Diamond$ сводится к сложению векторов, а операция: — к умножению вектора на скаляр, то оператор $\mathcal{O}_G\{\cdot\}$ может быть представлен в виде цепочки операторов, называемой гомоморфным фильтром

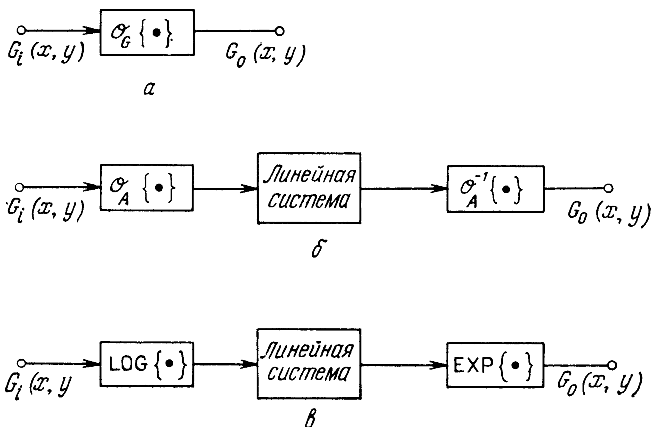


Рис. 1.8.1. Обобщенные линейные системы.

а — обобщенная система; *б* — представление обобщенной системы в виде гомоморфного фильтра; *в* — мультипликативный гомоморфный фильтр.

(рис. 1.8.1). Первый оператор $\mathcal{O}_A\{\cdot\}$ превращает операции $\Diamond$ и : в сложение векторов и умножение вектора на скаляр:

$$\mathcal{O}_A\{F_1(x, y) \Diamond F_2(x, y)\} = \mathcal{O}_A\{F_1(x, y)\} + \mathcal{O}_A\{F_2(x, y)\} \tag{1.8.3а}$$

и

$$\mathcal{O}_A\{k : F(x, y)\} = k\mathcal{O}_A\{F(x, y)\}. \tag{1.8.3б}$$

Вторая ступень гомоморфного фильтра — обычная линейная система. Третья ступень — оператор $\mathcal{O}_A^{-1}$, который является обратным относительно первого оператора, т. е.

$$\mathcal{O}_A^{-1}\{\mathcal{O}_A\{F(x, y)\}\} = F(x, y). \tag{1.8.4}$$

Рис. 1.8.1, *в* иллюстрирует частный случай гомоморфного фильтра для мультипликативной системы [5], в которой функция $G(x, y)$ получается в результате перемножения функций $F_1(x, y)$ и $F_2(x, y)$, т. е.

$$G(x, y) = F_1(x, y) \diamond F_2(x, y) \equiv F_1(x, y) F_2(x, y). \quad (1.8.5)$$

Прологарифмировав обе части равенства (1.8.5), получим сумму логарифмов функций $F_1(x, y)$ и $F_2(x, y)$:

$$\log \{G(x, y)\} = \log \{F_1(x, y)\} + \log \{F_2(x, y)\}. \quad (1.8.6)$$

Функция $\log G(x, y)$ преобразуется некоторой линейной системой, а затем посредством экспоненциального преобразования возвращается в пространство исходных изображений. Операция обобщенного умножения вектора на скаляр определяется как возведение в степень

$$G(x, y) = k : F(x, y) \equiv [F(x, y)]^k. \quad (1.8.7)$$

Логарифмирование этого равенства дает

$$\log \{G(x, y)\} = k \log \{F(x, y)\}. \quad (1.8.8)$$

Применение гомоморфной фильтрации при восстановлении изображений рассмотрено в гл. 15.

1.9. ВЕРОЯТНОСТНОЕ ОПИСАНИЕ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ [1]

Часто бывает удобно рассматривать изображение как реализацию случайного процесса. Введем порождающую изображения непрерывную случайную функцию $F(x, y, t)$ трех переменных — пространственных координат x, y и времени t.

Случайный процесс $F(x, y, t)$ полностью описывается совместной плотностью вероятности

$$p\{F_1, F_2, \ldots, F_J;\ x_1, y_1, t_1, x_2, y_2, t_2, \ldots, x_J, y_J, t_J\}$$

для J значений функции $F_j(x_j, y_j, t_j)$ в точках отсчета x_j, y_j, t_j. Совместные плотности вероятности высокого порядка для изображений обычно не известны, и их в общем случае трудно моделировать. Для плотности вероятности первого порядка $p(F; x, y, t)$ иногда удается подобрать удачную модель из физических соображений или на основе измеренных гистограмм. Например, плотность вероятности первого порядка случайного шума в элек-

[1]) Представленный в этом разделе материал предназначен читателю, знакомому с теорией вероятностей, случайными величинами и случайными процессами. Эти сведения можно найти в книгах [2, 6, 7]. Главная цель, которая преследовалась в этом разделе, — ввести соответствующие обозначения и рассмотреть вероятностные модели изображений.

тронных преобразователях изображений хорошо моделируется гауссовой плотностью:

$$p\{F;\ x,\ y,\ t\} = = [2\pi\sigma_F^2(x,\ y,\ t)]^{-1/2} \exp\left\{-\frac{[F(x,\ y),\ t-\eta_F(x,\ y,\ t)]^2}{2\sigma_F^2(x,\ y,\ t)}\right\}, \qquad (1.9.1)$$

где параметры $\eta_F(x,\ y,\ t)$ и $\sigma_F^2(x,\ y,\ t)$ есть среднее и дисперсия шума. Гауссова плотность может также с приемлемой точностью служить моделью плотности вероятности коэффициентов унитарных преобразований изображений. Плотность вероятности яркости должна быть односторонней, так как яркость принимает только положительные значения. В качестве моделей плотности вероятности яркости применяются плотность распределения вероятностей Рэлея

$$p\{F;\ x,\ y,\ t\} = \frac{F(x,\ y,\ t)}{\alpha^2} \exp\left\{-\frac{[F(x,\ y,\ t)]^2}{2\alpha^2}\right\}, \qquad (1.9.2а)$$

плотность логарифмически нормального распределения

$$p\{F;\ x,\ y,\ t\} = [2\pi F^2(x,\ y,\ t)\,\sigma_F^2(x,\ y,\ t)]^{-1/2} \times \times \exp\left\{-\frac{[\log[F(x,\ y,\ t)]-\eta_F(x,\ y,\ t)]^2}{2\sigma_F^2(x,\ y,\ t)}\right\} \qquad (1.9.2б)$$

и плотность экспоненциального распределения

$$p\{F;\ x,\ y,\ t\} = \alpha \exp\{-\alpha\,|F(x,\ y,\ t)|\}. \qquad (1.9.2в)$$

Эти плотности определены при $F \geqslant 0$, причем α — постоянная. Двусторонняя экспоненциальная, или лапласова, плотность

$$p\{F;\ x,\ y,\ t\} = (\alpha/2) \exp\{-\alpha\,|F(x,\ y,\ t)|\}, \qquad (1.9.3)$$

где α — постоянная, часто используется как модель плотности вероятности разностей отсчетов функции, описывающей изображение. Наконец, плотность равномерного распределения

$$p\{F;\ x,\ y,\ t\} = 1/2\pi, \quad -\pi \leqslant F \leqslant \pi, \qquad (1.9.4)$$

есть обычная модель для флуктуаций фазы случайного процесса. Для описания случайного процесса можно использовать также условные плотности вероятности. Условная плотность вероятности значения функции $F(x,\ y,\ t)$ в точке $(x_1,\ y_1,\ t_1)$ при заданном значении этой функции в точке $(x_2,\ y_2,\ t_2)$ определяется как

$$p\{F_1;\ x_1\ y_1,\ t_1\,|\,F_2;\ x_2,\ y_2,\ t_2\} = = \frac{p\{F_1,\ F_2;\ x_1,\ y_1,\ t_1,\ x_2,\ y_2,\ t_2\}}{p\{F_2,\ x_2,\ y_2,\ t_2\}}. \qquad (1.9.5)$$

Аналогично определяются условные плотности более высокого порядка.

Другой способ описания случайного процесса состоит в вычислении средних по ансамблю. Первый момент, или среднее значение функции $F(x, y, t)$, равен

$$\eta_F(x, y, t) = E\{F(x, y, t)\} = \int_{-\infty}^{\infty} F(x, y, t)\, p\{F; x, y, t)\, dF. \quad (1.9.6)$$

Второй момент, или автокорреляционная функция, определяется как

$$R(x_1, y_1, t_1; x_2, y_2, t_2) = E\{F(x_1, y_1, t_1) F^*(x_2, y_2, t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x_1, y_1, t_1) F^*(x_2, y_2, t_2)\, p\{F_1, F_2; x_1, y_1, t_1, x_2, y_2, t_2\}\, dF_1\, dF_2. \quad (1.9.7)$$

Автоковариационная функция изображения определяется как

$$K(x_1, y_1, t_1; x_2, y_2, t_2) = E\{[F(x_1, y_1, t_1) - \eta_F(x_1, y_1, t_1)] \times [F^*(x_2, y_2, t_2) - \eta_F^*(x_2, y_2, t_2)]\}. \quad (1.9.8а)$$

Автоковариационная и автокорреляционная функции связаны соотношением

$$K(x_1, y_1, t_1; x_2, y_2, t_2) = R(x_1, y_1, t_1; x_2, y_2, t_2) - \eta_F(x_1, y_1, t_1)\, \eta_F^*(x_2, y_2, t_2). \quad (1.9.8б)$$

Наконец, дисперсия процесса $F(x, y, t)$ есть

$$\sigma_F^2(x, y, t) = K(x, y, t; x, y, t). \quad (1.9.9)$$

Случайный процесс, порождающий изображения, называется стационарным в строгом смысле, если его моменты не зависят от переноса начала координат в пространстве или времени. Процесс называется стационарным в широком смысле, если он имеет постоянную среднюю яркость, а его автокорреляционная функция зависит от разностей координат $x_1 - x_2$, $y_1 - y_2$, $t_1 - t_2$, но не от самих координат. Для стационарного процесса $F(x, y, t)$

$$E\{F(x, y, t)\} = \eta_F \quad (1.9.10а)$$

и

$$R(x_1, y_1, t_1; x_2, y_2, t_2) = R(x_1 - x_2; y_1 - y_2; t_1 - t_2). \quad (1.9.10б)$$

Выражение для автокорреляционной функции можно записать в виде

$$R(\tau_x, \tau_y, \tau_t) = E\{F(x + \tau_x, y + \tau_y, t + \tau_t) F^*(x, y, t)\}. \quad (1.9.11)$$

Так как

$$R(-\tau_x, -\tau_y, -\tau_t) = R^*(\tau_x, \tau_y, \tau_t), \quad (1.9.12)$$

для действительной функции F автокорреляционная функция является действительной и четной. Энергетический спектр стационарного изображения по определению есть результат трехмерного преобразования Фурье его автокорреляционной функции:

$$\mathscr{W}(\omega_x, \omega_y, \omega_t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(\tau_x, \tau_y, \tau_t) \exp\{-i[\omega_x\tau_x + \omega_y\tau_y + \omega_t\tau_t]\} d\tau_x d\tau_y d\tau_t. \quad (1.9.13)$$

Во многих изображающих системах пространственные и временные процессы формирования изображений разделяются. В этом случае стационарную автокорреляционную функцию можно записать как

$$R(\tau_x, \tau_y, \tau_t) = R_{xy}(\tau_x, \tau_y) R_t(\tau_t). \quad (1.9.14)$$

Часто для упрощения вычислений пространственную автокорреляционную функцию представляют в виде произведения автокорреляционных функций для каждой пространственной переменной:

$$R_{xy}(\tau_x, \tau_y) = R_x(\tau_x) R_y(\tau_y). \quad (1.9.15)$$

В изображениях объектов, созданных человеком, часто встречаются горизонтальные и вертикальные структуры, поэтому аппроксимация автокорреляционной функции произведением (1.9.15) оказывается вполне приемлемой. В изображениях естественных сцен обычно нет преобладающих направлений корреляции. Пространственная автокорреляционная функция таких изображений близка к функции с вращательной симметрией и не является поэтому разделимой.

Часто моделью изображения служат реализации двумерного марковского процесса первого порядка. Автоковариационная функция такого процесса имеет вид

$$K_{x,y}(\tau_x, \tau_y) = K \exp\{-\sqrt{\alpha_x^2\tau_x^2 + \alpha_y^2\tau_y^2}\}, \quad (1.9.16)$$

где K, α_x и α_y — масштабные множители. Соответствующий энергетический спектр равен

$$\mathscr{W}(\omega_x, \omega_y) = (1/\alpha_x\alpha_y)\left[\frac{2K}{1 + [\omega_x^2/\alpha_x^2 + \omega_y^2/\alpha_y^2]}\right]. \quad (1.9.17)$$

Часто делается упрощающее предположение, что автоковариационная функция марковского процесса может быть представлена в виде

$$K_{x,y}(\tau_x, \tau_y) = K \exp\{-\alpha_x|\tau_x| - \alpha_y|\tau_y|\}. \quad (1.9.18)$$

Энергетический спектр этого процесса есть

$$\mathscr{W}(\omega_x, \omega_y) = (4\alpha_x\alpha_y K)/[(\alpha_x^2 + \omega_x^2)(\alpha_y^2 + \omega_y^2)]. \quad (1.9.19)$$

При детерминированном описании изображений были определены средние по пространству и времени. При статистическом описании также определяется среднее по ансамблю. Возникает вопрос: как связаны друг с другом пространственно-временные средние и средние по ансамблю? Ответ состоит в том, что для некоторых случайных процессов, называемых эргодическими, пространственно-временные средние и средние по ансамблю равны. Очень трудно доказать эргодичность случайного процесса в общем случае. Обычно достаточно определить эргодичность второго порядка, при которой моменты первого и второго порядка, полученные пространственно-временным усреднением, равны соответствующим моментам при усреднении по ансамблю.

1.10. ПРЕОБРАЗОВАНИЕ СЛУЧАЙНЫХ ИЗОБРАЖЕНИЙ В ИЗОБРАЖАЮЩЕЙ СИСТЕМЕ

Часто известны плотность вероятности или моменты случайного поля, представляющего изображение, на входе некоторой системы и требуется определить соответствующие характеристики для изображения на выходе этой системы. Если передаточная функция системы представлена в виде алгебраического выражения, то можно найти плотность вероятности на выходе посредством преобразования плотности вероятности на входе. Пусть, например, функции, описывающие изображения на входе и выходе системы, связаны следующим образом:

$$G(x, y, t) = \mathcal{O}_F\{F(x, y, t)\}, \tag{1.10.1}$$

где $\mathcal{O}_F\{\cdot\}$ — монотонный оператор, действующий на функцию $F(x, y, t)$. Тогда плотность вероятности на выходе есть

$$p\{G;\ x, y, t\} = \frac{p\{F(x, y, t)\}}{|d\mathcal{O}_F\{F(x, y)\}/dF|}. \tag{1.10.2}$$

Преобразование вида (1.10.2) можно распространить на плотности вероятности высокого порядка, однако часто полученные соотношения оказываются слишком громоздкими.

Мометны случайного поля на выходе системы могут быть получены непосредственно, если известна его плотность вероятности, или в некоторых случаях косвенным путем, исходя из свойств оператора системы. Если, например, этот оператор является линейным, то среднее значение поля $G(x, y, t)$ есть

$$E\{G(x, y, t)\} = E\{\mathcal{O}_L\{F(x, y, t)\}\} = \mathcal{O}_L\{E\{F(x, y, t)\}\}. \tag{1.10.3}$$

Можно показать, что если оператор системы линеен, а случайное поле на входе системы стационарно в строгом смысле, то случайное поле на выходе также стационарно в строгом смысле. Если же поле на входе стационарно в широком смысле, то в том же смысле стационарно поле на выходе.

Рассмотрим линейную пространственно-инвариантную систему, изображение на выходе которой представляется интегралом свертки:

$$G(x, y, t) = \iiint\limits_{-\infty}^{\infty} F(x-\alpha, y-\beta, t-\gamma) H(\alpha, \beta, \gamma)\, d\alpha\, d\beta\, d\gamma, \quad (1.10.4)$$

где $H(x, y, t)$ — импульсный отклик системы. Среднее значение функции $G(x, y, t)$ есть

$$E\{G(x, y, t)\} = \iiint\limits_{-\infty}^{\infty} E\{F(x-\alpha, y-\beta, t-\gamma)\} H(\alpha, \beta, \gamma)\, d\alpha\, d\beta\, d\gamma. \quad (1.10.5)$$

Если случайное поле на входе стационарно, то среднее значение η_F постоянно и может быть вынесено из под интеграла. Тогда

$$E\{G(x, y, t)\} = \eta_F \iiint\limits_{-\infty}^{\infty} H(\alpha, \beta, \gamma)\, d\alpha\, d\beta\, d\gamma = \eta_F \mathscr{H}(0, 0, 0), \quad (1.10.6)$$

где $\mathscr{H}(0, 0, 0)$ есть частотная характеристика линейной системы в начале координат пространственно-временной частотной области. Таким же способом легко показать, что автокорреляционные функции входного и выходного полей связаны соотношением

$$R_G(\tau_x, \tau_y, \tau_t) = R_F(\tau_x, \tau_y, \tau_t) * H(\tau_x, \tau_y, \tau_t) * H^*(-\tau_x, -\tau_y, \tau_t). \quad (1.10.7)$$

Выполнив преобразование Фурье обеих частей равенства (1.10.7) и используя теорему о преобразовании Фурье свертки, получим соотношение между энергетическими спектрами входного и выходного полей

$$\mathscr{W}_G(\omega_x, \omega_y, \omega_t) = \mathscr{W}_F(\omega_x, \omega_y, \omega_t) \mathscr{H}(\omega_x, \omega_y, \omega_t) \mathscr{H}^*(\omega_x, \omega_y, \omega_t), \quad (1.10.8а)$$

$$\mathscr{W}_G(\omega_x, \omega_y, \omega_t) = \mathscr{W}_F(\omega_x, \omega_y, \omega_t) |\mathscr{H}(\omega_x, \omega_y, \omega_t)|^2. \quad (1.10.8б)$$

Этот результат будет полезен при анализе действия шумов в изображающих системах.

ЛИТЕРАТУРА

1. Goodman J. W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Гудмен Дж., Введение в фурье-оптику. — М.: Мир, 1970.]
2. Papoulis A., Systems and Transforms with Applications in Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Папулис А., Теория систем и преобразований в оптике. — М.: Мир, 1971.]

3. Prewitt J. M. S., Object Enhancement and Extraction, in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970.
4. Oppenheim A. V., Generalized Superposition, *Information and Control*, **11**, 528—536 (November, 1967).
5. Oppenheim A. V., Schafer R. W., Stockham T. G., Jr., Nonlinear Filtering of Multiplied and Convolved Signals, *Proc. IEEE*, **56**, 8, 1264—1291 (August 1968). [Имеется перевод: Оппенхейм, Шефер, Стокхэм, Нелинейная фильтрация сигналов, представленных в виде произведения и свертки. — ТИИЭР, 1968, т. 56, № 8.]
6. Papoulis A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
7. Thomas J. B., An Introduction to Applied Probability Theory and Random Processes, Wiley, New York, 1971.

Глава 2

ПСИХОФИЗИЧЕСКИЕ СВОЙСТВА ЗРЕНИЯ

На выходе изображающих систем обычно создается фотоснимок или изображение на экране, которые рассматриваются человеком. Поэтому очевидно, что для эффективной разработки таких систем очень полезно понимать механизм человеческого зрения. Зная этот механизм, можно строить функциональные модели зрительной системы. Такие модели очень важны при разработке систем обработки изображений и создании критериев верности воспроизведения и дешифрируемости изображения.

2.1. ВОСПРИЯТИЕ СВЕТА

Согласно словарю Вебстера [1], свет есть «лучистая энергия, которая, действуя на органы зрения, дает им возможность выполнять зрительные функции». Физические свойства света изучены довольно хорошо, но механизм взаимодействия света с органами зрения еще не вполне понятен. Как известно, свет есть форма электромагнитного излучения, лежащего в относительно узкой области спектра в диапазоне длин волн от 350 до 780 нм. Характеристикой источника света является интенсивность излучения заданной длины волны. В зрительную систему человека попадает свет либо от самосветящегося источника, либо отраженный от некоторого предмета, либо прошедший через него. Пусть $E(\lambda)$ есть спектральная плотность излучения первичного источника света, а $t(\lambda)$ и $r(\lambda)$ — коэффициент пропускания и коэффициент отражения соответственно. Тогда для предмета, пропускающего свет, спектральная плотность излучения наблюдаемого света будет

$$C(\lambda) = t(\lambda) E(\lambda), \tag{2.1.1}$$

а для предмета, отражающего свет,

$$C(\lambda) = r(\lambda) E(\lambda). \tag{2.1.2}$$

На рис. 2.1.1 представлены графики спектральной плотности излучения некоторых обычных источников света, применяемых в изображающих системах: солнечного, вольфрамовой лампы накаливания, светоизлучающего диода, ртутной дуговой лампы и

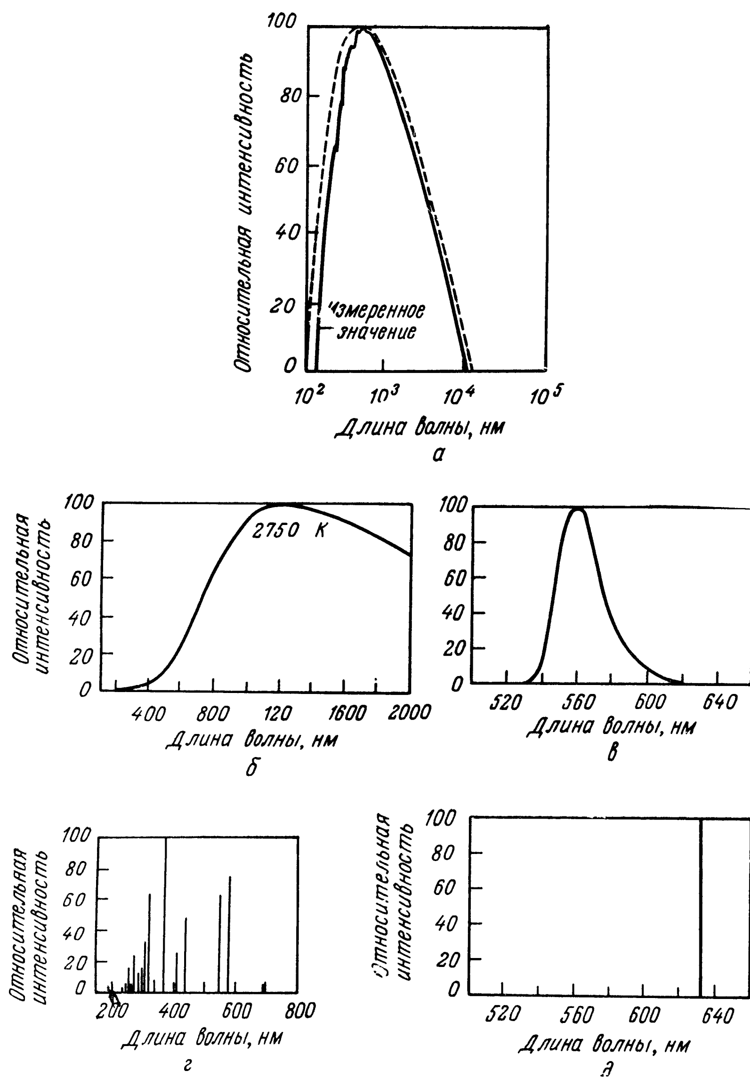


Рис. 2.1.1. Спектральная плотность излучения обычных источников света [2]. *а* — солнечный свет; *б* — вольфрамовая лампа накаливания; *в* — светоизлучающий диод; *г* — ртутная дуговая лампа; *д* — гелий-неоновый лазер.

гелий-неонового лазера. Свет от этих источников воспринимается зрителями по-разному. Солнечный свет кажется очень ярким и желтовато-белым, а свет лампы накаливания — менее ярким и довольно желтым. Светоизлучающий диод дает тусклый зеленый свет, а ртутная лампа — очень яркий, голубовато-белый. Наконец, лазер создает чрезвычайно яркий световой пучок чистого красного цвета. Эти наблюдения порождают много вопросов. Например, полностью ли описывает спектральная плотность излучения особенности восприятия этого излучения? Как будет видно, получен только частичный ответ на этот вопрос.

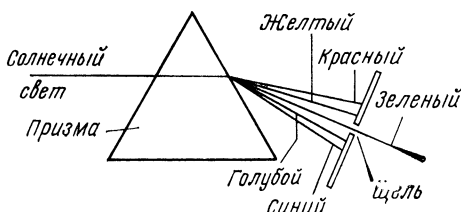


Рис. 2.1.2. Преломление света в призме.

Существуют три основные характеристики ощущения света — светлота, цветовой тон и насыщенность. Эти характеристики рассматриваются ниже.

Если имеются два источника света с одинаковой формой спектральной плотности излучения, то источник с большей интенсивностью излучения воспринимается как более яркий. Известно, однако, много примеров, когда предмет с одинаковой интенсивностью излучения всех точек не воспринимается как имеющий равномерную светлоту. Следовательно, интенсивность света не является адекватной количественной мерой светлоты.

Признак, которым отличается, например, красный свет от зеленого, называется цветовым тоном. С помощью призмы и щели, показанных на рис. 2.1.2, можно получать почти монохроматический свет различного цвета. Ясно, однако, что длина волны не является адекватной мерой цвета, так как в природе встречаются такие цвета, которые не наблюдаются в радуге, создаваемой призмой. Отсутствует, например, пурпурный цвет, который может быть получен смешением в равных количествах узкополосных красного и синего света. Если два источника света с одинаковыми спектральными плотностями наблюдать в одинаковых условиях, их цветовой тон будет одинаковым. Однако можно взять два таких источника света с разными спектральными плотностями, которые будут восприниматься как имеющие одинаковый цветовой тон (такие два излучения называются метамерической парой.)

Третья характеристика ощущения света — его насыщенность. Этот признак позволяет отличать спектральный цвет от пастельного блеклого цвета такого же цветового тона. По существу насыщенность описывает «белизну» цвета. Можно измерять насыщенность в процентах, сравнивая данный свет со спектральным цветом

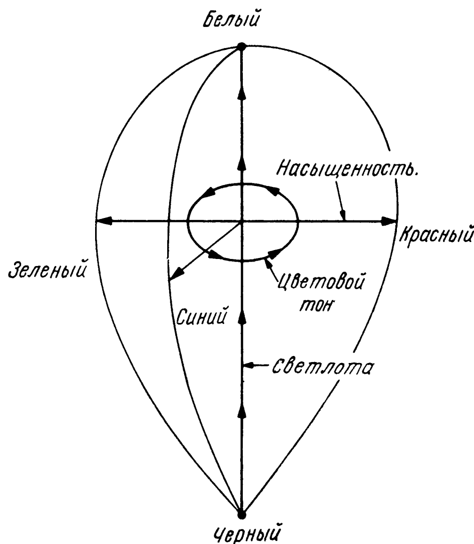


Рис. 2.1.3. Субъективные характеристики цвета.

на графике цветностей, изображенном на рис. 3.4.3. Однако обычно насыщенность не используют как количественную характеристику.

Для классификации цветов удобно рассматривать их как точки некоторого цветового пространства (рис. 2.1.3). Но для того чтобы можно было делать количественные выводы, расстояние между двумя точками в цветовом пространстве должно соответствовать субъективно воспринимаемой разнице между представляемыми цветами, где бы ни находилась эта пара цветов. Такие цветовые пространства будут рассматриваться далее.

2.2. ГЛАЗ

Наиболее естественным подходом к разработке модели зрительной системы человека было бы, видимо, проведение физиологического анализа глаза, нервных путей от глаза к мозгу и тех отделов мозга, которые связаны со зрительным восприятием. К сожалению, мы не в состоянии решить эту задачу из-за того,

что зрительная система состоит из огромного числа очень маленьких элементов. Однако в результате физиологических исследований глаза уже получено много данных, полезных для создания модели зрительной системы [4—7].

На рис. 2.2.1 показан поперечный разрез человеческого глазного яблока. Передняя прозрачная наружная оболочка глаза называется роговицей. Остальная часть наружной оболочки — склера — состоит из плотных волокон. Следующий слой — сосудистая оболочка, содержащая капилляры, которые снабжают

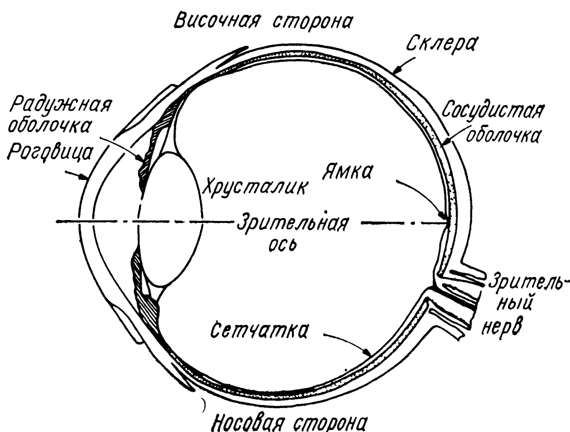


Рис. 2.2.1. Поперечный разрез глаза.

глаз кровью. Внутри сосудистой оболочки находится сетчатка с рецепторами двух типов — палочками и колбочками. Соединенные с сетчаткой нервные волокна выходят из глазного яблока в виде пучка — зрительного нерва. Свет, попавший в глаз через роговицу, фокусируется на поверхности сетчатки хрусталиком, форма которого под действием специальной мышцы меняется для обеспечения хорошей фокусировки предметов, находящихся на различном расстоянии от глаза. Радужная оболочка действует как диафрагма, изменяя количество света, проходящее в глаз.

Палочки — это длинные тонкие рецепторы, а колбочки — более короткие и толстые. Эти рецепторы действуют различным образом. Палочки более чувствительны к свету, чем колбочки. При малой освещенности палочки обеспечивают реакцию зрительной системы (ночное зрение). Колбочки функционируют при большой освещенности, их реакция представляет дневное зрение. На рис. 2.2.2 приведены относительные чувствительности палочек и колбочек в зависимости от длины волны воспринимаемого света [7, 8]. Глаз содержит около 6,5 млн. колбочек и 100 млн. пало-

чек. Распределение палочек и колбочек по сетчатке показано на рис. 2.2.3 [4]. Наибольшая плотность колбочек приходится на небольшую область сетчатки, называемую центральной ямкой которая находится вблизи выхода из глаза зрительного нерва. Это область наиболее резкого дневного зрения. В окрестности зрительного нерва нет ни палочек, ни колбочек — это слепое пятно глаза.

В последние годы установлено, что имеются три основных типа колбочек сетчатки [9, 10]. Эти колбочки имеют различные спек-

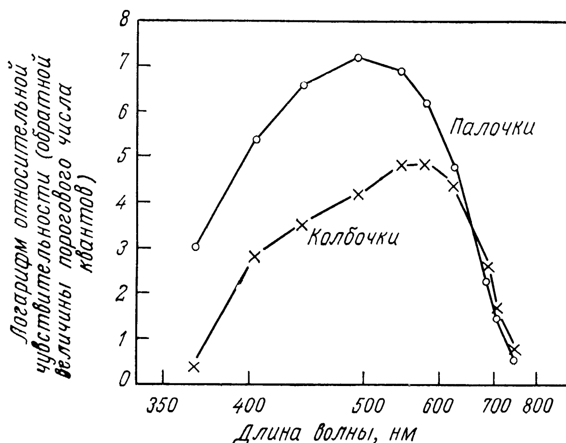


Рис. 2.2.2. Чувствительность палочек и колбочек по данным Валда [7, 8].

тральные характеристики поглощения света с максимумом в красной, зеленой и синей областях оптического спектра. На рис. 2.2.4 приведены спектральные кривые поглощения пигментов сетчатки [10]. Следует отметить две особенности этих кривых: во-первых, относительно низкую чувствительность колбочек типа α, воспринимающих в основном синий свет, и, во-вторых, значительное перекрытие кривых. Существование трех типов колбочек служит физиологической основой для трехцветной теории цветового зрения.

Когда свет возбуждает палочку или колбочку, возникает фотохимический переходный процесс, в результате которого создается нервный импульс. Механизм распространения нервных импульсов в зрительной системе в настоящее время полностью не выяснен. Известно, что зрительный нерв содержит около 800 000 нервных волокон. Сетчатка имеет свыше 100 000 000 рецепторов. Поскольку фотохимические процессы в сетчатке и механизм распространения нервных импульсов в глазе недостаточно изучены, нельзя дать полного описания зрительных процессов. Приходится удовлет-

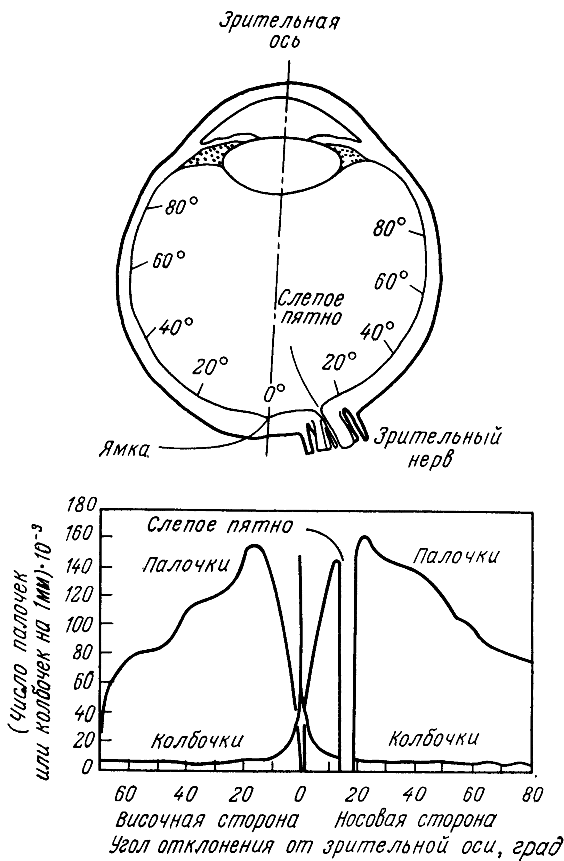


Рис. 2.2.3. Распределение палочек и колбочек по сетчатке [4].

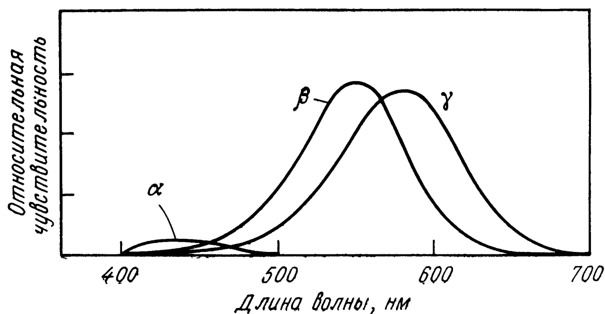


Рис. 2.2.4. Спектральные кривые поглощения пигментов рецепторов сетчатки [10].

вориться разработкой моделей, которые описывают и, будем надеяться, предскажут реакцию зрительной системы человека на те или иные изображения. В следующем разделе описываются некоторые явления, которые должны учитываться при моделировании зрительной системы.

2.3. ЗРИТЕЛЬНЫЕ ЯВЛЕНИЯ

Описанные ниже зрительные явления взаимосвязаны, одни — в меньшей степени, другие — в большей. Однако для упрощения изложения, а также потому, что некоторые детали явлений не ясны, эти явления рассматриваются вне связи друг с другом.

Контрастная чувствительность

Известно, что реакция глаза на изменение освещения является нелинейной. Пусть пятно света, имеющее интенсивность $I + \Delta I$, окружено фоном интенсивности I (рис. 2.3.1, *а*). Нужно найти

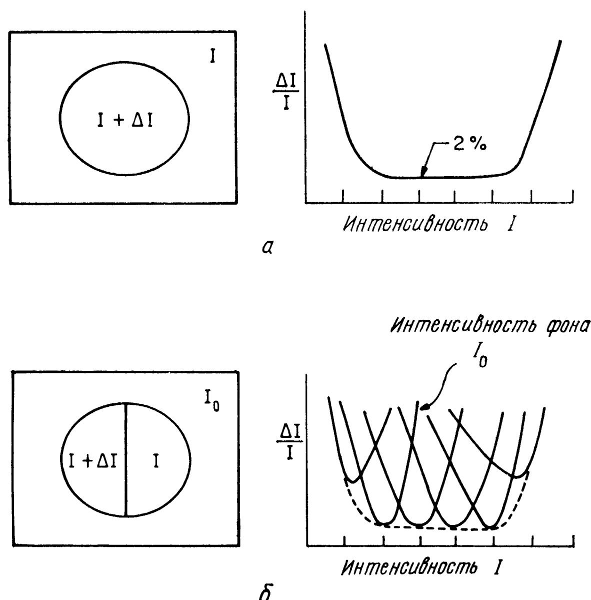


Рис. 2.3.1. Измерения контрастной чувствительности.
а — без фона; *б* — с фоном.

зависимость от I минимального приращения ΔI, которое замечается наблюдателем. Установлено [11], что в широком диапазоне интенсивностей отношение $\Delta I/I$, называемое отношением Вебера, имеет почти одинаковую величину — около 0,02. Этот результат

не имеет места для очень малых или очень больших интенсивностей, как видно из рис. 2.3.1, *а* [12]. Более того, контрастная чувствительность зависит от интенсивности окружающего фона. Рассмотрим два пятна света с интенсивностями I и $I + \Delta I$ (рис. 2.3.1, *б*), окруженные фоном интенсивности I_0. Графики зависимости отношения $\Delta I/I$ от I для разных интенсивностей фона I_0 приведены на рис. 2.3.1, *б* (справа). Оказалось, что диапазон интенсивностей, при котором отношение $\Delta I/I$ остается постоянным, значительно уже, чем на рис. 2.3.1, *а*. Огибающая точек минимума кривых рис. 2.3.1, *б* совпадает с кривой рис. 2.3.1, *а*. Однако диапазон интенсивностей I для почти постоянного отношения $\Delta I/I$ при фиксированной интенсивности фона I_0 сравним с динамическим диапазоном большинства электронных изображающих систем.

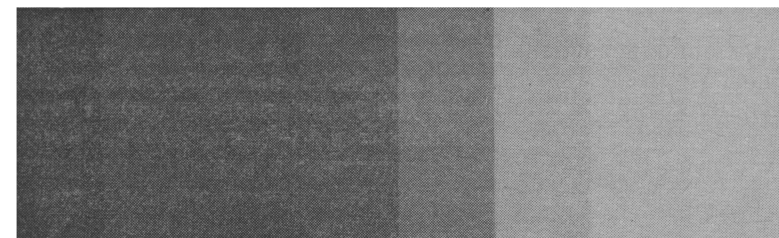
Поскольку дифференциал логарифма интенсивности равен

$$d\,\{\log(I)\} = dI/I, \tag{2.3.1}$$

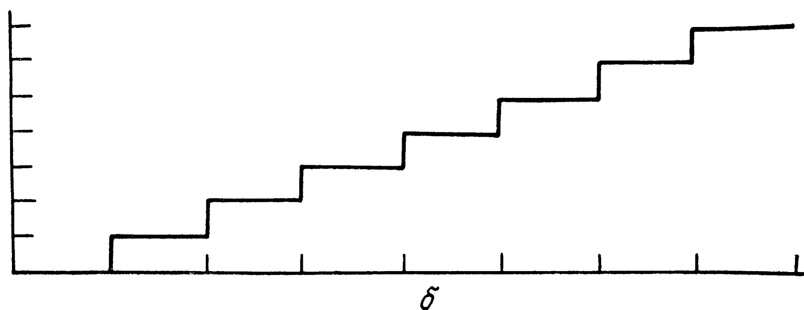
то одинаковые изменения логарифма интенсивности света можно связать с одинаковыми едва заметными изменениями интенсивности в диапазоне, в котором отношение $\Delta I/I$ остается постоянным. По этой причине во многих системах обработки изображений вычислительные операции проводятся с использованием логарифма интенсивности, а не интенсивности.

Полосы Маха

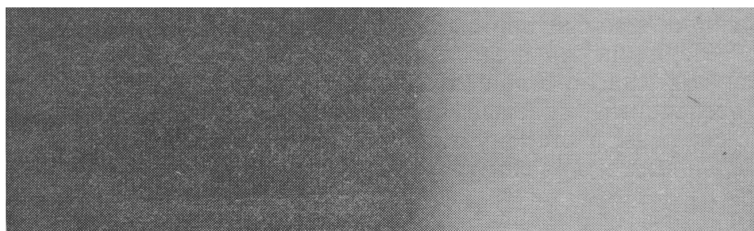
Рассмотрим ступенчатый оптический клин (набор полосок различной светлоты), показанный на рис. 2.3.2, *а*. Интенсивность света, отраженного от каждой полоски, одинакова по ее ширине и отличается на постоянную величину от интенсивности света, отраженного от соседних полосок. Однако правый край каждой полоски кажется темнее левого. Это явление называется эффектом полос Маха [13]. На рис. 2.3.2, *в* представлено изображение с распределением интенсивности, которое показано на рис. 2.3.2, *г*. На этой фотографии видны два столбика — яркий (B) и темный (D). Ни один из них нельзя предсказать исходя из распределения интенсивности. Полосы Маха, т. е. кажущиеся преувеличения изменений светлоты, можно объяснить исходя из пространственно-частотной характеристики глаза. Как мы увидим скоро, глаз обладает меньшей чувствительностью в области низких и высоких пространственных частот по сравнению с чувствительностью на средних частотах. Отсюда можно сделать вывод, что при создании систем обработки изображений можно в некоторой степени пожертвовать верностью воспроизведения контуров, так как глаз не очень чувствителен к высокочастотным перепадам светлоты.



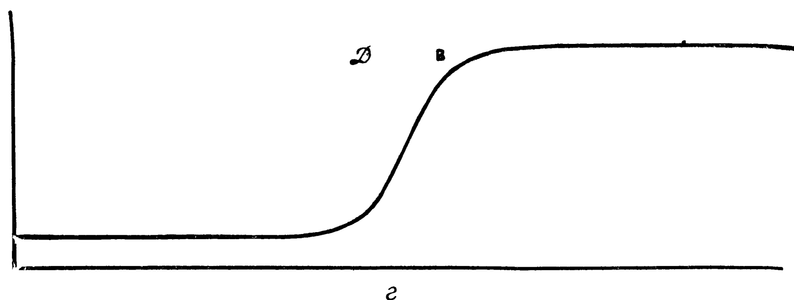
а



б



в



г

Рис. 2.3.2. Примеры полос Маха.

а — ступенчатый клин; *б* — распределение интенсивности клина; *в* — перепад интенсивности; *г* — распределение интенсивности для картины «в».

Одновременный контраст

Явление одновременного контраста [7] иллюстрируется фотографией, приведенной на рис. 2.3.3. В действительности все маленькие квадраты имеют одинаковую яркость, но из-за различной яркости фона кажется, что их яркость разная. Цветовой фон пятна света зависит также от цвета окружения. Белое пятно на черном квадрате кажется желтоватым, если вся фигура окружена синим фоном.

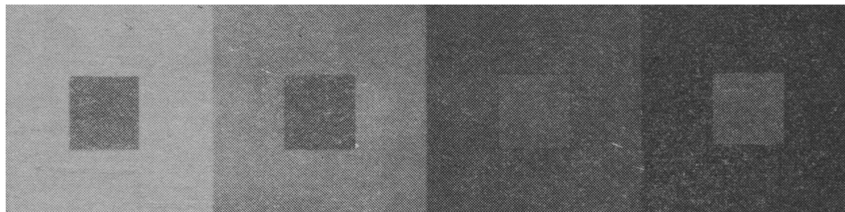


Рис. 2.3.3. Пример одновременного контраста.

Цветовая адаптация

Воспринимаемый цветовой фон зависит от адаптации зрителя [14]. Американский флаг, например, не сразу будет восприниматься как красно-бело-синий, если человек перед этим смотрел на интенсивный красный свет. Воспринимаемые цвета флага будут смещены в сторону голубого цвета, который является дополнительным к красному.

Цветовая слепота

Приблизительно 8% всех мужчин и 1% женщин в той или иной форме страдают цветовой слепотой [15, стр. 405]. Существуют разные степени цветовой слепоты. Некоторые люди, так называемые монохроматы, имеют только палочки или палочки и колбочки одного типа и поэтому им свойственно одноцветное зрение. Другие люди — дихроматы — имеют колбочки двух типов. И монохроматы, и дихроматы могут различать цвета в той степени, в которой они научились связывать различные цвета с разными предметами. Например, они могут знать, что темные розы — красные, а светлые — желтые. Но если покрасить красную розу в желтый цвет, сохранив величину коэффициента отражения, монохромат может назвать ее красной. Дихроматы также неспособны точно определять цветовой тон.

Субъективные цвета [16]

В 1826 г. французский монах Бенидикт Прево заметил, что если периодически освещать белый лист бумаги, помещая его в узкий луч света в затемненной комнате, то на этом листе можно наблюдать цветные полоски. Это было первое зарегистрированное наблюдение субъективных цветов. Фехнер в 1838 г. наблюдал субъективные цвета, появляющиеся на вращающемся диске,

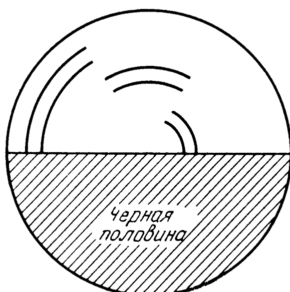


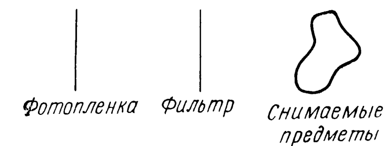
Рис. 2.3.4. Диск Бенхэма.

выкрашенном наполовину в черный, наполовину в белый цвет. В 1894 г. Бенхэм изобрел волчок, на верхней поверхности которого был узор, показанный на рис. 2.3.4. Когда волчок вращается против часовой стрелки, внешнее кольцо кажется красным, среднее — зеленым, а внутреннее — синим. Вращение в противоположную сторону приводит к тому, что цвета внешнего и внутреннего колец меняются местами. Эти и связанные с ними явления частично объясняются характером временной реакции зрительной системы человека на вспышки света.

Опыты Лэнда

В начале 50-х годов Лэнд [17] провел интересные опыты, показывающие возможность воспроизведения цвета с помощью двух основных цветов. Один из этих опытов иллюстрируется на рис. 2.3.5. Некоторая сцена фотографируется дважды: за фильтрами с полосами пропускания ~600—700 и ~500—600 нм. Изготавливаются два одноцветных диапозитива T_1 и T_2. Затем эти диапозитивы проектируются на один экран. Первый диапозитив освещается красным светом (длина волны 600—700 нм), а второй — белым. Наблюдается интересный эффект: когда оба изображения совмещены, исходная сцена воспроизводится почти в естественных цветах (плохо воспроизводятся только пурпурные

цвета). Если изображения слегка не совпадают, воспроизведение цветов резко ухудшается. Это показывает, что воспроизведение цветов зависит от содержания сцены, а не есть чисто локальное явление (т. е. цвет данной точки зависит от цвета ее окружения).



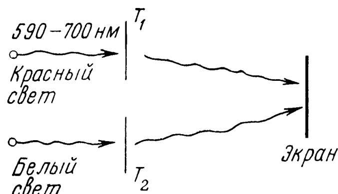


Рис. 2.3.5. Опыты Лэнда.

2.4. МОДЕЛЬ ОДНОЦВЕТНОГО ЗРЕНИЯ

При использовании одного из современных методов разработки оптических систем оптическую систему рассматривают как линейную систему относительно интенсивности света, которую можно описать с помощью двумерной частотной характеристики [18]. Исследуем линейную оптическую систему, показанную на рис. 2.4.1. Пространственное распределение интенсивности света на входе системы получено после прохождения пучка света постоянной интенсивности через диапозитив, прозрачность которого меняется по синусоидальному закону. Вследствие линейности

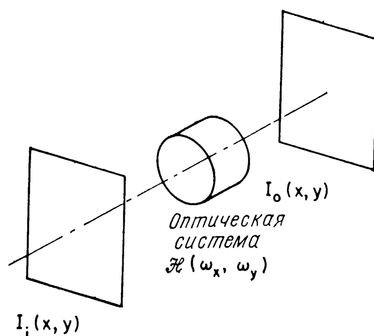


Рис. 2.4.1. Линейная оптическая система.

системы распределение интенсивности света на выходе также будет синусоидальным, но, возможно, отличным по амплитуде или фазе от распределения на входе. Изменяя пространственную частоту (число периодов изменения интенсивности на единицу длины) распределения интенсивности света на входе и измеряя амплитуду и фазу распределения интенсивности на выходе,

можно в принципе получить частотную характеристику оптической системы.

Пусть $\mathscr{H}(\omega_x, \omega_y)$ — оптическая частотная характеристика двумерной линейной системы, а $\omega_x = 2\pi/T_x$, $\omega_y = 2\pi/T_y$ — угловые пространственные частоты (T_x, T_y — периоды изменения интенсивности вдоль координатных осей x и y). Пусть далее $I_i(x, y)$ обозначает распределение интенсивности света на входе в предметной плоскости, а $I_o(x, y)$ — распределение интенсивности света на выходе (в плоскости изображения). Тогда пространственно-частотные спектры распределений интенсивностей на входе и выходе по определению будут равны

$$\mathscr{I}_i(\omega_x, \omega_y) = \iint\limits_{-\infty}^{\infty} I_i(x, y) \exp\{-i(\omega_x x + \omega_y y)\}\, dx\, dy \qquad (2.4.1)$$

и

$$\mathscr{I}_o(\omega_x, \omega_y) = \iint\limits_{-\infty}^{\infty} I_o(x, y) \exp\{-i(\omega_x x + \omega_y y)\}\, dx\, dy. \qquad (2.4.2)$$

Эти спектры связаны следующим соотношением:

$$\mathscr{I}_o(\omega_x, \omega_y) = \mathscr{H}(\omega_x, \omega_y)\, \mathscr{I}_i(\omega_x, \omega_y). \qquad (2.4.3)$$

Распределение интенсивности света на выходе может быть получено обратным преобразованием фурье-спектра (2.4.2):

$$I_o(x, y) = (1/4\pi^2) \iint\limits_{-\infty}^{\infty} \mathscr{I}_o(\omega_x, \omega_y) \exp\{i(\omega_x x + \omega_y y)\}\, d\omega_x\, d\omega_y. \qquad (2.4.4)$$

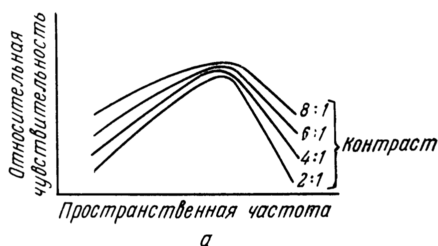
Во многих случаях при разработке оптических систем интересуются только отношением амплитуд интенсивностей на входе и выходе. Это отношение

$$|\mathscr{I}_o(\omega_x, \omega_y)|/|\mathscr{I}_i(\omega_x, \omega_y)| = |\mathscr{H}(\omega_x, \omega_y)| \qquad (2.4.5)$$

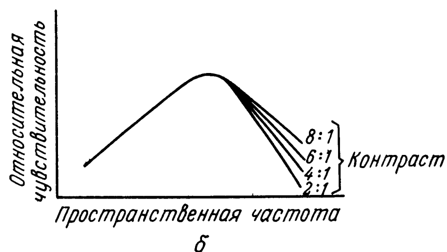
называется частотно-контрастной характеристикой (ЧКХ) оптической системы.

Много работ было посвящено вопросу использования понятия линейной системы для описания свойств зрительной системы человека [19—25]. Типичный эксперимент для проверки справедливости линейной модели зрительной системы заключается в следующем. Наблюдателю показывают две решетки с синусоидальным изменением коэффициента пропускания. Одна из них — опорная, с постоянным контрастом [1] и заданной пространственной частотой, вторая — испытательная, с переменным контрастом и частотой, отличающейся от частоты опорной решетки. Кон-

[1] Контрастом пространственно-периодического распределения интенсивности называется отношение максимальной интенсивности к минимальной.



а



б

Рис. 2.4.2. Результаты измерения ЧКХ зрительной системы человека.

а — синусоидальная решетка: *б* — экспоненциально-синусоидальная решетка.

Рис. 2.4.3. Сложная синусоидальная решетка.

траст испытательной решетки изменяется до тех пор, пока яркость светлых и темных штрихов обеих решеток будет казаться одинаковой. Применяя набор испытательных решеток с разными пространственными частотами, можно построить график ЧКХ зрительной системы человека. На рис. 2.4.2, *а* приведены графики ЧКХ для разных контрастов входного сигнала. Представление о форме ЧКХ можно также получить, рассматривая сложную синусоидальную решетку, показанную на рис. 2.4.3. В этой решетке пространственная частота возрастает по одной координате, а контраст — по другой. Огибающая различимых штрихов в общем близка по форме к кривым рис. 2.4.2, *а* [24].

На рис. 2.4.2, *а* показано, что форма ЧКХ зависит от контраста распределения интенсивности на входе. Кроме того, она меняется при вращении испытательной решетки вокруг оптической оси глаза. Таким образом, зрительная система человека нелинейна и анизотропна (не обладает инвариантностью относительно поворота).

Было сделано предположение, что нелинейная реакция глаза на изменения интенсивности имеет логарифмический характер. Нелинейное преобразование осуществляется на первом этапе обработки визуальной информации в зрительной системе. Преоб-

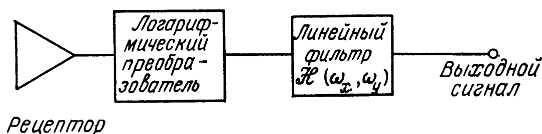


Рис. 2.4.4. Логарифмическая модель одноцветного зрения.

разованию подвергаются сигналы рецепторов (палочек и колбочек) до того, как они вступают во взаимодействие друг с другом. На рис. 2.4.4 приведена простейшая логарифмическая модель глаза для одноцветного зрения. Если реакция глаза на величину входной интенсивности действительно имеет логарифмический характер, то, применив экспоненциальное преобразование интенсивности синусоидальной решетки, можно «линеаризовать» зрительную систему. ЧКХ, полученная в результате измерения с использованием экспоненциально-синусоидальной решетки (рис. 2.4.2, *б*), хорошо аппроксимируется одной кривой на низких и средних пространственных частотах. На рис. 2.4.5 приведены результаты измерения ЧКХ, полученные Дэвидсоном [26] для экспоненциально-синусоидального испытательного сигнала. Высокочастотная часть этой кривой получена экстраполированием входного сигнала среднего контраста.

Логарифмически-линейная модель глаза (рис. 2.4.3) позволяет удовлетворительно предсказать реакцию зрительной системы

в широком диапазоне интенсивностей. Однако на высоких пространственных частотах и очень низких или очень высоких интенсивностях наблюдаемая реакция отличается от предсказанной на основе этой модели. Для того чтобы получить более точную модель, необходимо детальнее рассмотреть физические процессы в зрительной системе.

Нелинейная реакция палочек и колбочек на изменения интенсивности все еще служит предметом активных исследований. Были высказаны гипотезы о том, что механизм этой нелинейности объясняется химической активностью, электрическими эффектами и нервной обратной связью. Основная логарифмическая модель предполагает, что

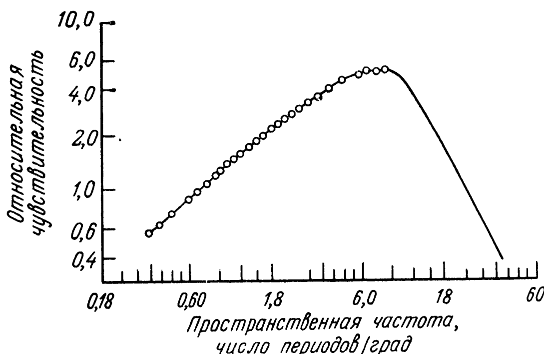


Рис. 2.4.5. Результаты измерений ЧКХ зрительной системы человека с помощью экспоненциально-синусоидальных решеток [26].

$$I_o(x, y) = K_1 \log [K_2 + K_3 I_i(x, y)], \qquad (2.4.6)$$

где K_1, K_2 и K_3 — постоянные, а $I_i(x, y)$ и $I_o(x, y)$ — входной и выходной сигналы. В другой модели [7, стр. 253] предполагается следующая зависимость выходного сигнала от входного:

$$I_o(x, y) = [K_1 I_i(x, y)]/[K_2 + I_i(x, y)], \qquad (2.4.7)$$

где K_1, K_2 — постоянные. Маннос и Сакрисон [27] изучали различные нелинейности, используемые в аналитических мерах верности воспроизведения изображения. Они установили, что степенная нелинейная зависимость, когда

$$I_o(x, y) = [I_i(x, y)]^s, \qquad (2.4.8)$$

где s — постоянная, обеспечивает хорошее согласие между вычисленной верностью и субъективной оценкой качества изображения. (Подробно результаты этой работы рассматриваются в гл. 7.) Подбором постоянных, входящих в выражения (2.4.6)—

(2.4.8), соответствующие кривые могут быть достаточно хорошо сближены в диапазоне средних интенсивностей.

Пространственно-частотные характеристики глаза объясняются частично оптическими и частично нервными механизмами. Как оптический инструмент глаз имеет ограниченную разрешающую способность из-за конечных размеров апертуры линзы, оптических аберраций и конечных размеров палочек и колбочек. Эти эффекты могут быть представлены в модели фильтром нижних

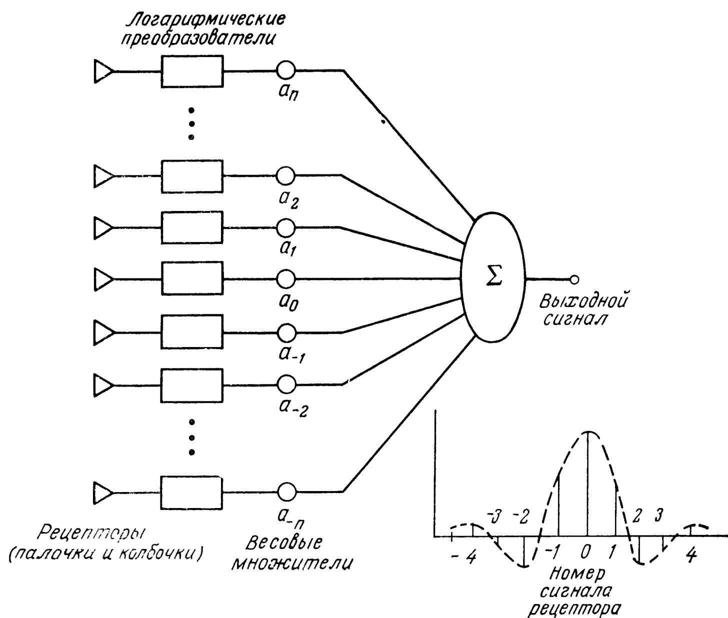


Рис. 2.4.6. Схема латерального торможения.

пространственных частот, включенным между рецептором и нелинейным элементом. Наиболее существенный вклад в частотную характеристику глаза вносит механизм латерального торможения [28]. Схема этого механизма приведена на рис. 2.4.6. Нервный сигнал есть взвешенная сумма сигналов от многих соседних рецепторов. Некоторые рецепторы вносят тормозящий (т. е. отрицательный) вклад в эту сумму. Весовые множители представляют собой по существу импульсный отклик части зрительной системы, находящейся после сетчатки. Двумерное преобразование Фурье этого импульсного отклика дает соответствующую частотную характеристику.

Восприятие светового импульса, попавшего в глаз, происходит с задержкой. После окончания действия импульса ощущение

света затягивается на короткое время. Эти явления, возникающие из-за ненулевой длительности временно́й реакции зрительной системы, можно моделировать линейным временны́м фильтром.

На рис. 2.4.7 показана модель одноцветного зрения, учитывающая результаты, которые обсуждались выше. В этой модели выходной сигнал рецептора подается на вход линейного фильтра нижних частот, соответствующего оптике глаза. Далее включен элемент с монотонно нелинейной характеристикой, моделиру-

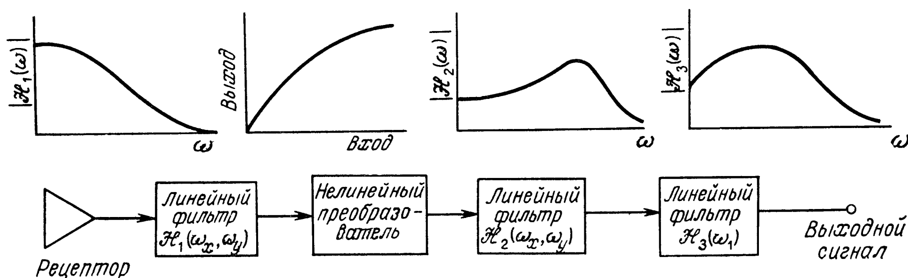


Рис. 2.4.7. Расширенная модель одноцветного зрения.

ющий нелинейную реакцию рецепторов на изменение интенсивности. Процесс латерального торможения учитывается введением полосового линейного пространственного фильтра. Эффекты временно́й фильтрации моделируются последним звеном — линейной системой. В работе [29] тщательно исследована эта модель и получены характеристики ее отдельных звеньев, при которых с большой точностью моделируется полная реакция зрительной системы. При удачном выборе масштабных коэффициентов рассмотренная модель одноцветного зрения будет, по-видимому, достаточно точной для большинства задач обработки изображений. Даже упрощенная логарифмическая модель (рис. 2.4.4), вероятно, будет во многих случаях удовлетворительной.

2.5. МОДЕЛЬ ЦВЕТОВОГО ЗРЕНИЯ

Начиная с опытов Ньютона и Максвелла [30—33], было предложено множество теорий, объясняющих цветовое зрение человека. В классической трехцветной модели цветового зрения, разработанной Юнгом в 1802 г. [32], предполагается, что глаз обладает тремя видами элементов, чувствительных в разных зонах оптического спектра. Интересно отметить, что до 1960 г. не было прямых физиологических подтверждений существования трех различных типов чувствительных элементов сетчатки [9, 10].

На рис. 2.5.1 показана схема модели цветового зрения, предложенная Фреем [34]. В этой модели три рецептора со спектраль-

ными чувствительностями $s_1(\lambda)$, $s_2(\lambda)$ и $s_3(\lambda)$, соответствующими трем пигментам колбочек сетчатки, создают сигналы

$$e_1 = \int C(\lambda)\, s_1(\lambda)\, d\lambda, \qquad (2.5.1а)$$

$$e_2 = \int C(\lambda)\, s_2(\lambda)\, d\lambda, \qquad (2.5.1б)$$

$$e_3 = \int C(\lambda)\, s_3(\lambda)\, d\lambda, \qquad (2.5.1в)$$

где $C(\lambda)$ — спектральная плотность энергии источника падающего света. Три сигнала e_1, e_2, e_3 подвергаются затем логарифмическому преобразованию и объединяются так, чтобы получились сигналы

$$d_1 = \log e_1, \qquad (2.5.2а)$$

$$d_2 = \log e_2 - \log e_1 = \log (e_2/e_1), \qquad (2.5.2б)$$

$$d_3 = \log e_3 - \log e_1 = \log (e_3/e_1). \qquad (2.5.2в)$$

Эти сигналы проходят через линейные фильтры с частотными характеристиками $\mathscr{H}_1(\omega_x, \omega_y)$, $\mathscr{H}_2(\omega_x, \omega_y)$ и $\mathscr{H}_3(\omega_x, \omega_y)$; получаемые на выходе сигналы g_1, g_2, g_3 определяют восприятие цветов в мозге.

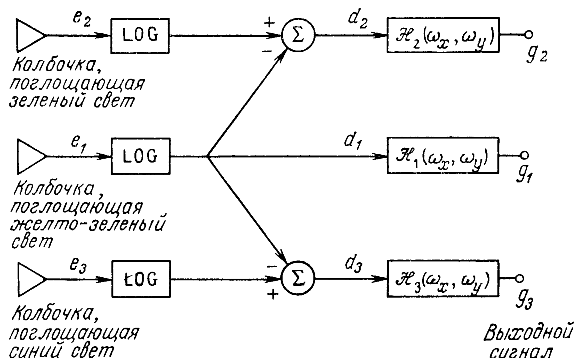


Рис. 2.5.1. Модель цветового зрения.

В модели рис. 2.5.1 сигналы d_2 и d_3 характеризуют цветность, а сигнал d_1 пропорционален яркости. Оказалось, что эта модель позволяет очень точно предсказать многие явления цветового зрения и хорошо согласуется с основными законами колориметрии. Известно, например, что если спектральную плотность энергии источника света умножить на постоянную (одинаковую для всех длин волн), то цветовой тон и насыщенность, описываемые координатами цветности, останутся неизменными в широком диапазоне изменения интенсивности света. Выражения (2.5.1) и (2.5.2)

показывают, что сигналы цветности d_2 и d_3 в таком случае не изменяются, а сигнал яркости изменяется по логарифмическому закону. Другие особенности этой модели описаны Фреем [34].

Как уже отмечалось, некоторые данные о спектральной чувствительности s_i (λ) трех типов колбочек сетчатки были получены при измерении поглощения света пигментами колбочек (см. рис. 2.2.4). Однако прямые физиологические измерения очень сложны и не могут быть выполнены с большой точностью. Косвенные оценки спектральной чувствительности колбочек были получены Конигом

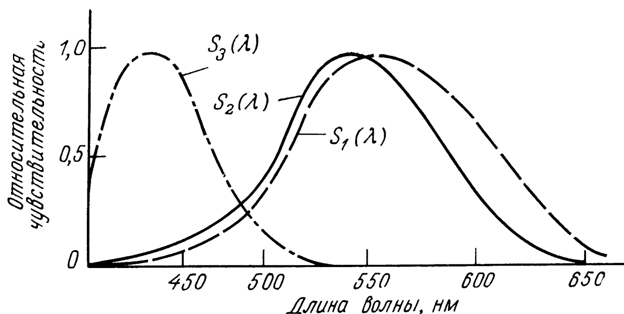


Рис. 2.5.2. Спектральные чувствительности колбочек по данным Конига [35].

и Бродхуном [35] при изучении аномалий цветового зрения. Джадд [36] на основе этих данных нашел линейное преобразование, позволяющее установить связь спектральных чувствительностей s_i (λ) с функциями сложения, найденными в колориметрических опытах. В результате были получены кривые, приведенные на рис. 2.5.2. Они унимодальны и строго положительны, как это и следует из физиологических представлений.

Аналогично модели одноцветного зрения (рис. 2.4.7) логарифмическая модель цветового зрения (рис. 2.5.1) может быть дополнена линейными фильтрами, которые включаются после рецепторов. Вместо логарифмической функции можно использовать нелинейную функцию общего вида. Заметим, что, не изменяя выходного сигнала, можно изменить порядок выполнения линейных операций суммирования и преобразования. Схема расширенной модели цветового зрения представлена на рис. 2.5.3. Можно ожидать, что пространственно-частотная характеристика яркостного канала, на выходе которого образуется сигнал g_1, будет подобна пространственно-частотной характеристике системы одноцветного зрения, которая рассматривалась в разд. 2.4. Результаты измерений частотной характеристики системы в окрашенном свете приведены на рис. 2.5.4. Как видно, частотная характеристика, измеренная при использовании окрашенного света, сдвинута

к низким пространственным частотам по сравнению с частотной характеристикой, измеренной с белым светом [37]. Латеральное торможение должно приводить к изгибу частотной характеристики

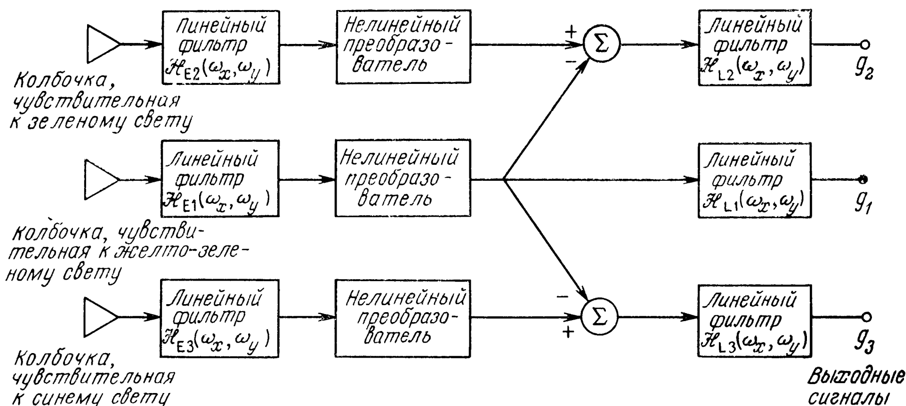


Рис. 2.5.3. Расширенная модель цветового зрения.

на низких частотах. Этот изгиб получается, вероятно, на более низких частотах, чем те, которые входили в исследуемый диапазон.

Воспринимаемый цвет — понятие относительное. Ощущение, вызываемое светом с данным спектральным распределением энергии, зависит от окружающего фона и адаптации зрителя. Человек

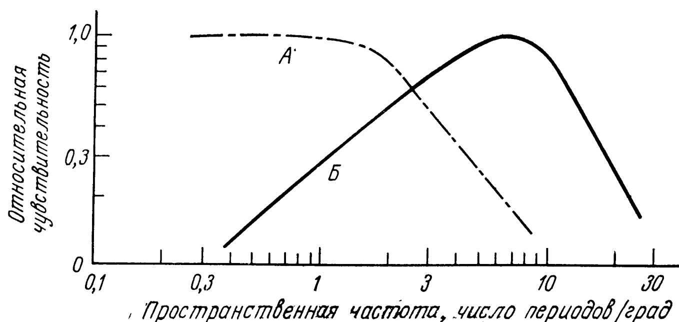


Рис. 2.5.4. Пространственно-частотные характеристики зрительной системы человека [26, 37].

А — для скрашенного света; *Б* — для белого света.

может очень хорошо адаптироваться к освещению сцены, используя опорный белый свет или общий цветовой баланс. Это свойство называется цветовой адаптацией.

В простейшей модели цветовой адаптации, предложенной фон Крисом [38; 15, стр. 435], в схему зрительной системы (рис. 2.5.3)

между колбочками и первым линейным фильтром включается блок автоматической регулировки усиления. Коэффициент усиления

$$a_i = \left[\int W(\lambda)\, s_i(\lambda)\, d\lambda \right]^{-1} \qquad (2.5.3)$$

для $i = 1, 2, 3$ устанавливается так, чтобы усиленный сигнал колбочки равнялся единице при рассматривании опорного белого света со спектральной плотностью $W(\lambda)$. Модель фон Криса привлекает разумностью и простотой. Однако эксперименты показали [15, стр. 438], что эта модель не полностью описывает эффект цветовой адаптации. Уоллис [39] высказал предположение о том, что цветовая адаптация может быть частично объяснена действием механизма торможения, в результате которого ослабляются медленно изменяющиеся компоненты яркостного поля. Этот механизм может моделироваться фильтрами с характеристиками $\mathcal{H}_{Li}(\omega_x, \omega_y)$ (рис. 2.5.3). Несомненно, оба механизма — регулировки усиления и торможения — обеспечивают цветовую адаптацию. Требуются дальнейший анализ и эксперименты, чтобы правильно объяснить это явление и построить его модель.

ЛИТЕРАТУРА

1. Webster's New Collegiate Dictionary, G. and C. Merriam Co., The Riverside Press, 1960.
2. Malitson H. H., The Solar Energy Spectrum, *Sky and Telescope*, **29**, 4, 162—165 (March 1965).
3. Munsell Book of Color, Munsell Color Co., 2441 North Calvert St., Baltimore, Maryland.
4. Pirenne M. H., Vision and the Eye, 2nd ed., Associated Book Publishers, London, 1967.
5. Polyak S. L., The Retina, University of Chicago Press, Chicago, 1941.
6. Davson L. H., The Physiology of the Eye, McGraw-Hill (Blakiston), New York, 1949.
7. Cornsweet T. N., Visual Perception, Academic press, New York, 1970.
8. Wald G., Human Vision and the Spectrum, *Science*, **101**, 2635, 653—658 (June 1945).
9. Brown P. K., Wald G., Visual Pigment in Single Rods and Cones of the Human Retina, *Science*, **144**, 3614, 45—52 (April 1964).
10. Wald G., The Receptors for Human Color Vision, *Science*, **145**, 3636, 1007—1017 (September 1964).
11. Hecht S., The Visual Discrimination of Intensity and the Weber-Fechner Law, *J. Gen, Physiol.*, **7**, 241 (1924).
12. Stevens S. S., Handbook of Experimental Psychology, Wiley, New York, 1951.
13. Ratliff F., Mach Bands: Quantitative Studies on Neural Networks in the Retina, Holden-Day, San Francisco, 1965.
14. Brindley G. S., Afterimages, *Scientific American*, 209, 4, 84—93 (October 1963).
15. Wyszecki G., Stiles W. S., Color Science, Wiley, New York, 1967.
16. Cohen J., Gordan D., The Prevost-Fechner-Benham Subjective Colors, Psychophys. Bull., **46**, 2, 97—136 (March 1949).

17. Land E. H., Experiments in Color Vision, Scientific American, **200**, 5, 84—99 (May 1959).
18. Goodman J. W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Гудмен Дж., Введение в фурье-оптику. — М.: Мир, 1970.]
19. Campbell F. W., The Human Eye as an Optical Filter, *Proc. IEEE*, **56**, 6, 1009—1014 (June 1968).
20. Bryngdahl O., Characteristics of the Visual System: Psychophysical Measurement of the Response to Spatial Sine-Wave Stimuli in the Mesopic Region, *J. Opt. Soc. Am.*, **54**, 9, 1152—1160 (September 1964).
21. Lowry E. M., Depalma J. J., Sine Wave Response of the Visual System, I. The Mach Phenomenom, *J. Opt. Soc. Am.*, **51**, 7, 740—746 (July 1961).
22. Lowry E. M., Depalma J. J., Sine Wave Response of the Visual System, II. Sine Wave and Square Wave Contrast Sensitivity, *J. Opt. Soc. Am.*, 52, 3, 328—335 (March 1962).
23. Sachs M. B., Nachmias J., Robson J. G., Spatial Frequency Channels in Human Vision, *J. Opt. Soc. Am.*, **61**, 9, 1176—1186 (September 1971).
24. Stockham T. G., Jr., Image Processing in the Context of a Visual Model, *Proc. IEEE*, **60**, 7, 828—842 (July 1972). [Имеется перевод: Стокхэм, Обработка изображений в контексте модели зрения. — ТИИЭР, 1972, т. 60, № 7.]
25. Pearson D. E., A Realistic Model for Visual Communication Systems, *Proc. IEEE*, **55**, 3, 380—389 (March 1967). [Имеется перевод: Пирсон, Реалистическая модель системы передачи изображений. — ТИИЭР, 1967, т. 55, № 3.]
26. Davidson M. L., Perturbation Approach to Spatial Brightness Interaction in Human Vision, *J. Opt. Soc. Am.*, **58**, 9, 1300—1308 (September 1968).
27. Mannos J. L., Sakrison D. J., The Effects of a Visual Fidelity Criterion on the Encoding of Images, *IEEE Trans. Inf. Theory*, IT-**20**, 4, 525—536 (July 1974).
28. Ratliff F., Hartline H. K., Miller W. H., Spatial and Temporal Aspects of Retinal Inhibitory Interaction, *J. Opt. Soc. Am.*, **53**, 1, 110—120 (January 1963).
29. Hall C. F., Hall E. L., A Nonlinear Model for the Spatial Characteristics of the Human Visual System, *IEEE Trans. Syst., Man Cybern.*, SMC-**7**, 3, 161—170 (March 1977).
30. McCann J. J., Human Color Perception, in: Color Theory and Imaging Systems, Enyard R. A., Ed., Society of Photographic Scientists and Engineers, Washington D. C., 1973, pp. 1—23.
31. Newton I., Optiks, 4th ed., 1730; Dover Publications, New York, 1952. [Имеется перевод: Ньютон И., Оптика. — М.: Наука, 1954.]
32. Young T., Philosophical Transactions, **92**, 1802, 12—48.
33. Maxwell J. C., Scientific Papers of James Clerk Maxwell, Neven W. D., Ed., Dover Publications, New York, 1965.
34. Frei W., A New Model of Color Vision and Some Practical Limitations, University of Southern California, Image Processing Institute, USCEE Report 530, March 1974, pp. 128—143.
35. Konig A., Brodhun E., Experimentelle Untersuchungen uber die Psychophysische Fundamental in Bezug auf den Gesichtssinn, Zweite Mittlg. S. B. Preuss Akademic der Wissenschaften, 1889, 641.
36. Judd D. B., Standard Response Functions for Protanopic and Deuteranopic Vision, *J. Opt. Soc. Am.*, **35**, 3, 199—221 (March 1945).
37. Van der Horst C. J. C., de Weert C. M., Bouman M. A., Transfer of Spatial Chromaticity — Contrast at Threshold in the Human Eye, *J. Opt. Soc. Am.*, **57**, No. 10, 1260—1266 (October 1967).
38. von Kries J., Die Gesichtsempfindungen, Nagel's Handbuch d. Physiology d. Menchen, Vol. 3, 1904, p. 211.
39. Wallis R. H., Film Recording of Digital Color Images, University of Southern California Image Processing Institute Report No. 570, June 1975.

Глава 3

ФОТОМЕТРИЯ И КОЛОРИМЕТРИЯ

В гл. 2 зрение человека рассматривалось в качественном аспекте и была сделана попытка построить модели одноцветного и цветового зрения. Этим моделям можно придать количественный характер, если ввести меры величин, определяющих восприятие света человеком. Меры яркости изучаются в фотометрии, а цвета — в колориметрии. В данной главе рассматриваются вопросы, относящиеся к обеим этим дисциплинам, а также источники света

3.1. ИСТОЧНИКИ СВЕТА

Источник лучистой энергии описывается спектральной плотностью $C(\lambda)$ — скоростью испускания энергии в единичном интервале длин волн. Полная мощность излучения источника определяется интегралом спектральной плотности — лучистым потоком

$$P = \int_0^\infty C(\lambda)\, d\lambda. \tag{3.1.1}$$

Эта величина обычно измеряется в ваттах.

Тело, имеющее более высокую температуру, чем окружающая среда, излучает электромагнитную энергию. Чем выше температура, тем сильнее излучение. Идеальный тепловой излучатель называется абсолютно черным телом. Лучистый поток абсолютно черного тела при любом значении длины волны достигает максимума, возможного для заданной температуры. Спектральная плотность излучения определяется законом Планка [1]

$$C(\lambda) = \frac{C_1}{\lambda^5\,[\exp\{C_2/\lambda T - 1\}]}, \tag{3.1.2}$$

где λ — длина волны излучения, T — температура тела, C_1, C_2 — постоянные. На рис. 3.1.1, *а* показаны графики спектральной плотности для разных значений температуры. В видимой области электромагнитного спектра спектральная плотность из-

лучения абсолютно черного тела может быть аппроксимирована законом Вина [1]:

$$C(\lambda) = \frac{C_1}{\lambda^5 \exp\{C_2/\lambda T\}}. \qquad (3.1.3)$$

График функции Вина для видимой области спектра приведен на рис. 3.1.1, *б*.

Главнейшим источником света является, конечно, солнце. На рис. 2.1.1, *а* приведен график измеренного спектрального рас-

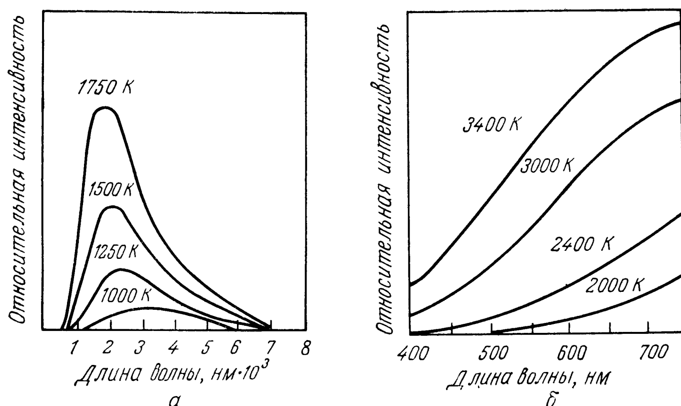


Рис. 3.1.1 Излучение абсолютно черного тела [1].
а — закон Планка; *б* — закон Вина.

пределения энергии солнечного света [2]. Пунктирная линия на этом рисунке, аппроксимирующая результаты измерений, представляет собой график спектральной плотности излучения абсолютно черного тела при $T = 6000$ K. Лампы накаливания также часто приближенно рассматривают как абсолютно черное тело при $T = 1500—3500$ K [3].

Международная комиссия по освещению (МКО) утвердила несколько стандартных источников света, спектральные плотности излучения которых графически представлены на рис. 3.1.2 [4]. Источник S_A есть лампа накаливания с вольфрамовой нитью. Источник S_B дает свет, близкий к прямому солнечному свету в диапазоне 400—700 нм. Свет источника S_C близок к свету при облачном небе. Источник E с равномерной спектральной плотностью излучения часто применяется при колориметрических расчетах.

В системах обработки изображений в качестве источников света часто используются люминофоры кинескопов. На рис. 3.1.3 представлены спектральные плотности излучения нескольких обычных люминофоров [5]. В одноцветных телевизионных прием-

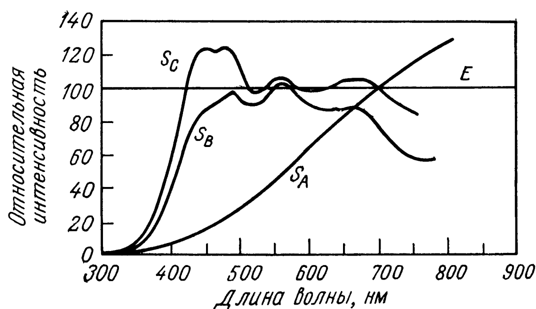


Рис. 3.1.2. Стандартные источники МКО [4].

Относительная энергия излучения

Цветовая температура 7000 К

Длина волны, нм

Люминофор сульфидного типа P4

Люминофор P16

Люминофор P24

Люминофор P22

Рис. 3.1.3. Спектральная плотность излучения люминофоров кинескопов [5].

никах обычно применяют люминофор P4, который обеспечивает воспроизведение довольно яркой голубовато-белой картинки. Люминофор P16 используется в видеодатчиках бегущего луча, так как имеет короткое послесвечение. Для видеодатчиков бегущего луча цветного телевидения выбирают люминофор P24, излучающий в относительно широкой спектральной полосе и имеющий довольно короткое послесвечение. В приемниках цветного телевидения используют кинескопы с красным, зеленым и синим люминофорами, расположенными в виде триад точек или полос. Люминофор P22 представляет собой типичную комбинацию трех люминофоров.

3.2. ФОТОМЕТРИЯ

Цель фотометрических измерений — количественно описать яркость света [6—8]. Связь между фотометрическими и энергетическими характеристиками излучения устанавливается функцией световой эффективности, показанной на рис. 3.2.1, *а* [9]. Эта кривая утверждена как стандарт МКО. Она определяет спектральную

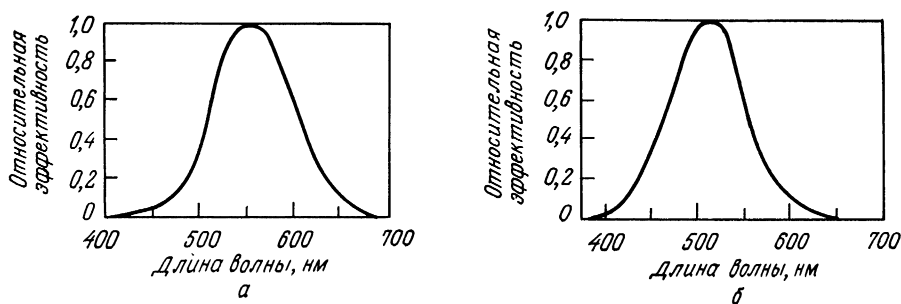


Рис. 3.2.1. Относительная спектральная световая эффективность [9].
а — для дневного зрения; *б* — для ночного зрения.

чувствительность зрительной системы некоторого типичного (стандартного) наблюдателя. В сущности эта кривая есть стандартизованный вариант кривой спектральной чувствительности колбочек, приведенной на рис. 2.2.2, для дневного зрения при относительно высоких уровнях освещенности. Стандартная функция световой эффективности для ночного зрения при относительно низких уровнях освещенности представлена на рис. 3.2.1, *б*. Большинство систем воспроизведения изображений разрабатывается на основе относительной световой эффективности для дневного зрения.

Субъективное ощущение яркости, вызываемое источником света со спектральной плотностью излучения $C(\lambda)$, определяется его световым потоком, который выражается как

$$F = K_m \int_0^\infty C(\lambda) V(\lambda) d\lambda, \qquad (3.2.1)$$

где $V(\lambda)$ — относительная световая эффективность, а K_m — постоянная. Если световой поток измерять в люменах, то $K_m = 685$ лм/Вт. Источник монохроматического света мощностью

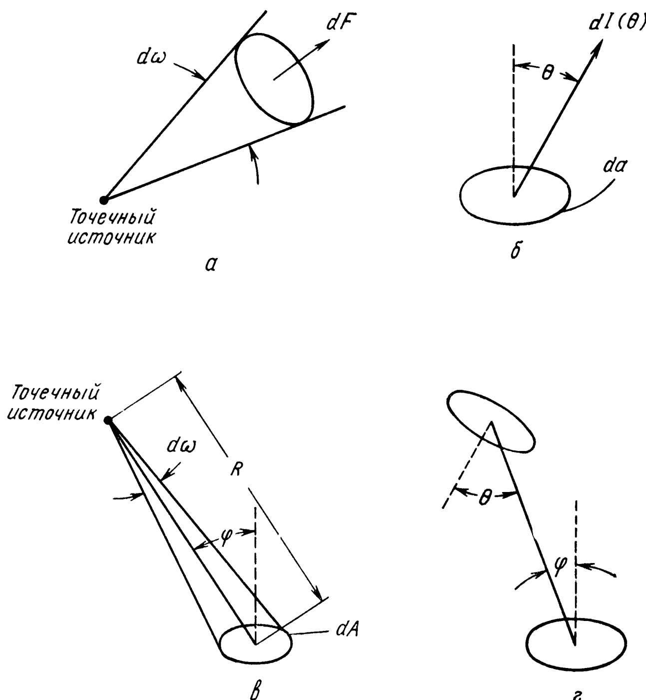


Рис. 3.2.2. К определению фотометрических величин.

а — излучение точечного источника; *б* — излучение элементарной площадки; *в* — освещенность, создаваемая точечным источником; *г* — освещенность, создаваемая протяженным источником.

1 Вт, длина волны которого соответствует максимуму кривой относительной световой эффективности (т. е. при $\lambda = 555$ нм), дает световой поток 685 лм.

Другой важной характеристикой источника света является пространственное распределение его излучения. Рассмотрим точечный источник, испускающий световой поток во всех направлениях (рис. 3.2.2, *а*). Отношение бесконечно малого светового по-

тока dF, излучаемого в бесконечно малый телесный угол $d\omega$, к величине этого угла называется силой света

$$I = dF/d\omega. \tag{3.2.2}$$

Эта величина измеряется в канделах. Протяженный источник характеризуется яркостью (рис. 3.2.2, *б*), равной по определению

$$B = dI(\theta)/(da\cos\theta), \tag{3.2.3}$$

где $dI(\theta)$ — бесконечно малая сила света, излучаемого с бесконечно малой площадки da в направлении, составляющем угол θ с нормалью к поверхности. Единицей яркости является кандела на 1 м^2, или нит.

Протяженный источник света является диффузным, если справедлив закон Ламберта

$$dI(\theta) = dI_N\cos\theta, \tag{3.2.4}$$

где dI_N — бесконечно малая сила света, излучаемого перпендикулярно поверхности. Для такого источника яркость

$$B = dI_N/da \tag{3.2.5}$$

не зависит от направления наблюдения. В изображающих системах освещенное матовое стекло часто используется как приближение к идеальному диффузному источнику света.

Яркость источника излучения есть мера светового потока, излучаемого с единицы площади его поверхности в заданном направлении в единичный телесный угол. Величина светового потока, падающего на наблюдаемую поверхность, определяется ее освещенностью. Пусть точечный источник испускает световой поток dF в телесный угол $d\omega$ под углом φ с нормалью к наблюдаемой поверхности (рис. 3.2.2, *в*). Угол $d\omega$ охватывает площадку dA на этой поверхности. Освещенность по определению равна

$$E = dF/dA. \tag{3.2.6}$$

Из соотношения (3.2.2) и рис. 3.2.2 следует, что

$$dF = I(\theta)\,d\omega = I(\theta)\,[(dA\cos\varphi)/R^2], \tag{3.2.7}$$

где R — расстояние от источника до поверхности. Следовательно, при освещении точечным источником освещенность определяется соотношением

$$E = [I(\theta)\cos\varphi]/R^2, \tag{3.2.8}$$

которое выражает закон обратных квадратов для излучения. Чтобы определить освещенность, создаваемую протяженным источником, его нужно представить в виде совокупности точечных источников, а затем сложить освещенности, получаемые от этих точечных источников. Обращаясь к рис. 3.2.2, *г* и выражению (3.2.8), получим, что бесконечно малая освещенность площадки dA

наблюдаемой поверхности, создаваемой площадкой da излучающей поверхности, есть

$$dE = [dI\,(\theta)\cos\varphi]/R^2. \qquad (3.2.9)$$

Из выражений (3.2.3) и (3.2.9) получаем

$$dE = (B\,da\cos\theta\cos\varphi)/R^2. \qquad (3.2.10)$$

Для вычисления освещенности E нужно интегрировать выражение (3.2.10) по поверхности освещающего предмета.

На основе приведенных выше теоретических положений можно получить два важных фотометрических результата для оптической системы. Рассмотрим обобщенную изображающую систему, показанную на рис. 3.2.3, *а*. Бесконечно малая площадка da поверхности диффузного предмета с яркостью B отображается в бесконечно малую площадку dA наблюдаемой поверхности, которая приобретает яркость $\tilde{B}$. Можно показать [8, стр. 189], что $\tilde{B} = B$, если коэффициенты отражения обеих поверхностей

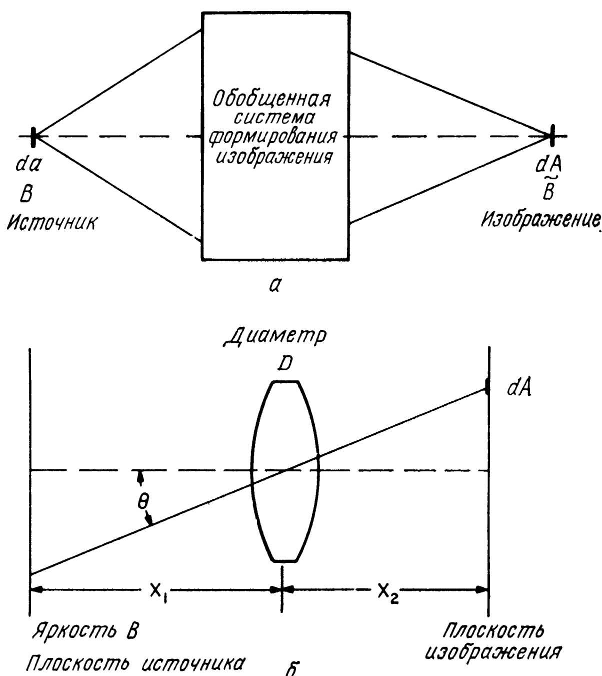


Рис. 3.2.3. Связь яркости источника с яркостью и освещенностью изображения. *а* — яркость источника и изображения; *б* — яркость источника и освещенность изображения.

одинаковы. Таким образом, изображающая система не изменяет яркость, если не принимать во внимание потерь из-за поглощения и рассеяния света. На рис. 3.2.3, *б* показано, как происходит отображение протяженного диффузного источника с яркостью B на наблюдаемую поверхность посредством простой тонкой линзы. Бесконечно малая освещенность на бесконечно малой площадке dA [8, стр. 190] равна

$$dE = \left(B\pi D^2/4x_2^2\right)\cos^4\theta, \tag{3.2.11}$$

где D — диаметр линзы, θ — угол пересечения главных лучей, проходящих через центр линзы, и x_2 — расстояние от центра линзы до точки наблюдения. Таким образом, освещенность точек плоскости наблюдения при удалении от оптической оси падает как четвертая степень угла между данными направлением и оптической осью. Заметим, что $\cos^4 30° = 0{,}56$. Дополнительные потери освещенности вызываются поглощением и рассеянием света в оптической системе. Кроме того, могут иметь место потери из-за виньетирования, когда полностью теряются косые лучи, составляющие большой угол с оптической осью.

В дальнейшем при рассмотрении систем обработки изображений, для того чтобы не входить в излишние детали, будут по возможности исключаться вопросы, связанные с фотометрией. Кроме того, как это обычно делается в литературе по обработке изображений, термин «яркость» будет использоваться для обозначения светового потока, силы света и собственно яркости.

3.3. УРАВНИВАНИЕ ЦВЕТОВ [10—14]

В основе трехцветной теории цветного зрения лежит возможность представления произвольного цвета сложением в нужной пропорции трех основных цветов. При аддитивном цветовоспроизведении, например в цветном телевидении, имеются три источника света — красного, зеленого и синего, которые проектируются на один участок поверхности, где и воспроизводится заданный цвет. При субтрактивном цветовоспроизведении, на котором основаны цветная фотография и полиграфия, белый свет приобретает нужный цвет, проходя последовательно через голубой, пурпурный и желтый фильтры.

3.3.1. АДДИТИВНОЕ УРАВНИВАНИЕ ЦВЕТОВ

Опыты с аддитивным уравниванием цветов проводятся по схеме, показанной на рис. 3.3.1. На поверхность идеального диффузного отражателя проектируется пятно света $[C]$ (рис. 3.3.1, *а*) с произвольной спектральной плотностью, приведенной на рис. 3.3.2, *а*. На ту же поверхность проектируются пятно опор-

ного белого цвета [W] со спектральной плотностью, показанной на рис. 3.3.2, *б*, а также перекрывающиеся пятна [P_1], [P_2], [P_3] трех основных цветов со спектральными плотностями, представ-

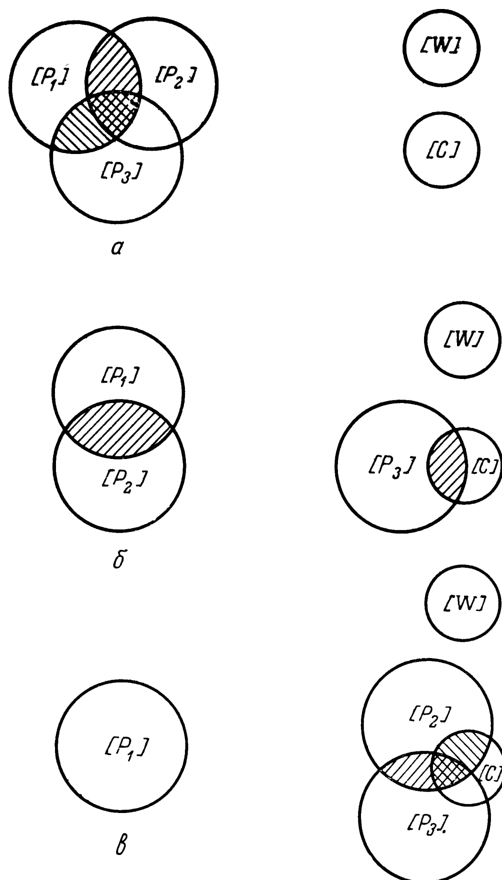


Рис. 3.3.1. Уравнивание цветов.
а — три основных цвета; *б* — два основных цвета; *в* — один основной цвет

ленными на рис. 3.3.2, *в*, *г*, *д*. Сначала интенсивности трех основных цветов устанавливаются так, чтобы общая часть пятен имела такие же светлоту, цветовой тон и насыщенность, как и опорный белый цвет [W]. Интенсивности A_1 (W), A_2 (W), A_3 (W) измеряют в каких-либо физических единицах, например в ваттах. Они являются величинами, уравнивающими опорный белый цвет. Затем интенсивности основных цветов устанавливают так, чтобы добиться уравнивания заданного цвета [C]. Если уравнивание

достигается, записывают интенсивности $A_1(C)$, $A_2(C)$, $A_3(C)$ и вычисляют нормированные величины

$$T_i(C) = A_i(C)/A_i(W), \quad i = 1, 2, 3, \tag{3.3.1}$$

которые называются координатами цвета.

Если данный цвет уравнять не удается, от схемы рис. 3.3.1, *а* переходят к схеме рис. 3.3.1, *б*. Один из основных цветов, например $[P_3]$, накладывается на пятно данного цвета, и интенсивности трех основных цветов подбирают до тех пор, пока не будет

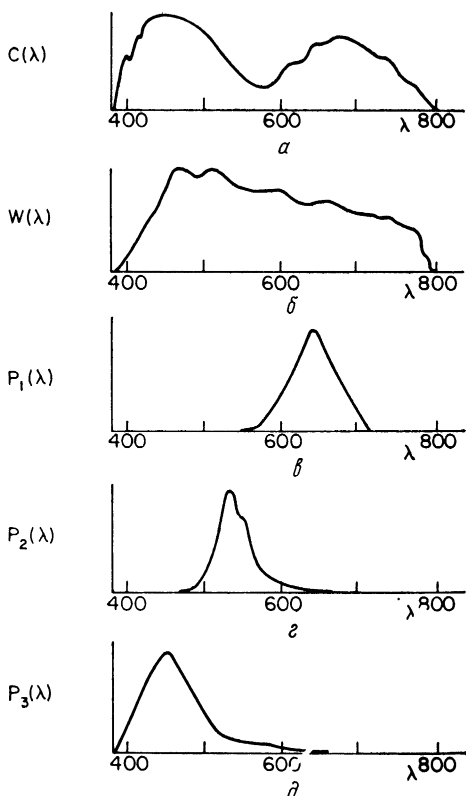


Рис. 3.3.2. Спектральная плотность энергии излучения.

а — произвольный цвет; *б* — опорный белый цвет; *в* — красный основной цвет; *г* — зеленый основной цвет; *д* — синий основной цвет

достигнуто уравнивание общей части пятен $[P_1]$ и $[P_2]$ и общей части пятен $[P_3]$ и $[C]$. Если такое уравнивание получено, то координаты цвета равны

$$T_1(C) = A_1(C)/A_1(W), \quad T_2(C) = A_2(C)/A_2(W),$$
$$T_3(C) = -A_3(C)/A_3(W). \tag{3.3.2}$$

В этом случае координата T_3 (C) отрицательная. Если при такой схеме добиться уравнивания не удается, нужно пробовать уравнивать сочетание [P_1] и [P_3] с сочетанием [P_2] и [C]. Если уравнивание будет достигнуто, координату T_2 (C) нужно взять со знаком минус. Если и эта конфигурация окажется несостоятельной, нужно попробовать уравнять [P_2] и [P_3] с [P_1] и [C].Правильное уравнивание в этом случае означает, что величина T_1 (C) отрицательная.

Наконец, в тех редких случаях, когда не удается добиться уравнивания ни по схеме рис. 3.3.1, *а*, ни по схеме рис. 3.3.1, *б*, нужно два основных цвета наложить на данный цвет [C] и пробовать уравнивать цвет общей частоты этих трех пятен с оставшимся основным цветом. Если уравнивание достигается для схемы рис. 3.3.1, *в*, то две координаты цвета принимают отрицательные значения:

$$T_1(C) = A_1(C)/A_1(W), \quad T_2(C) = -A_2(C)/A_2(W),$$
$$T_3(C) = -A_3(C)/A_3(W). \tag{3.3.3}$$

Если уравнивание по этой схеме не достигается, одна из двух оставшихся комбинаций приведет к успеху.

Описанный выше процесс реализует прямой метод количественного описания цветов. Он имеет два недостатка: является довольно громоздким и получаемые результаты зависят от индивидуальных свойств единственного наблюдателя. В разд. 3.4 подробно рассматривается стандартизованный метод количественного измерения цветов.

3.3.2. СУБТРАКТИВНОЕ УРАВНИВАНИЕ ЦВЕТОВ

Схема опытов по субтрактивному уравниванию цветов приведена на рис. 3.3.3. Свет со спектральной плотностью E (λ) последовательно проходит через три фильтра — голубой, пурпурный и желтый. Спектральная кривая поглощения каждого красителя зависит от его концентрации. Нужно заметить, что кривые поглощения используемых на практике красителей при изменении их концентрации меняются нелинейным образом.

На первом этапе процесса субтрактивного уравнивания концентрации красителей трех фильтров изменяются до тех пор, пока не будет достигнуто субъективное уравнивание опорного белого цвета [W]. Эти концентрации являются уравнивающими величинами A_1 (W), A_2 (W), A_3 (W). Затем концентрации изменяются так, чтобы обеспечить уравнивание желаемого цвета [C]. Найденные концентрации — уравнивающие величины A_1 (C), A_2 (C), A_3 (C) — используются при вычислении координат T_1 (C), T_2 (C), T_3 (C) согласно соотношению (3.3.1). Очевидно, что между аддитивным и субтрактивным уравниванием цветов принципиальной разницы нет. В субтрактивной системе желтый краситель дей-

ствует как переменный поглотитель синего света, т. е. управляет его интенсивностью. Таким же образом пурпурный фильтр управляет зеленым, а голубой — красным светом. Эти фильтры являются широкополосными и пропускают много света. Поэтому их удобнее использовать в субтрактивных системах, чем узко-

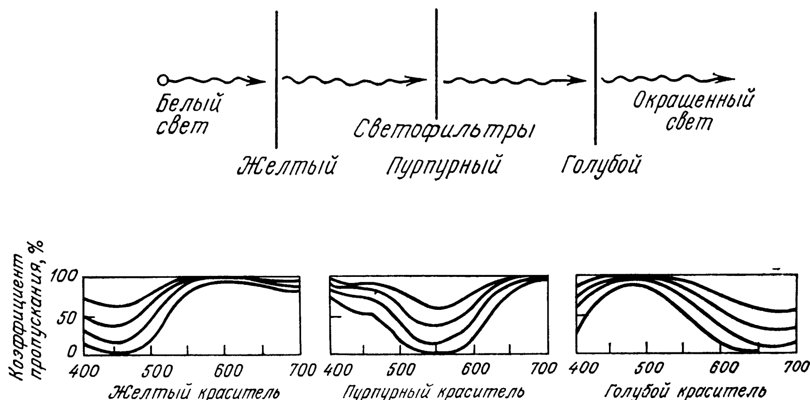


Рис. 3.3.3. Субтрактивное уравнивание цветов.

полосные красный, зеленый и синий фильтры. Обычные красители не обладают широким диапазоном изменения плотности. Чтобы повысить контраст воспроизводимого изображения, в полиграфии часто вводят четвертый фильтрующий слой — нейтрально-серый. Субтрактивное воспроизведение цветов с реальными красителями, обладающими перекрывающимися кривыми поглощения, рассматривается в гл. 16.

3.3.3. АКСИОМЫ УРАВНИВАНИЯ ЦВЕТОВ

Опыты по уравниванию цветов (как аддитивному, так и субтрактивному) проведены с большой точностью многими исследователями. Установлено, что полного уравнивания иногда не удается достигнуть при очень слабом или очень сильном освещении. Результаты уравнивания в некоторой степени зависят от цветового окружения. Тем не менее оказалось, что простые правила уравнивания цветов остаются справедливыми в широком диапазоне условий эксперимента

Грассман [15] ввел восемь аксиом, определяющих трехцветное уравнивание цветов, которые являются основой колориметрии:

1. Любой цвет может быть уравнен смесью не более чем трех цветов.

2. Уравнивание, достигнутое при данной интенсивности света, сохраняется в широком диапазоне интенсивностей.

3. Смесь цветов не может быть разделена человеческим глазом на отдельные компоненты.

4. Яркость смеси цветов равна сумме яркостей ее компонент.

5. Закон сложения. Если цвет $[M]$ эквивалентен цвету $[N]$, а цвет $[P]$ — цвету $[Q]$, то смесь цветов $[M]$ и $[P]$ эквивалентна смеси $[N]$ и $[Q]$ [1]:

$$[M] \triangleq [N],\ [P] \triangleq [Q] \Rightarrow [M] \mathbin{\stackrel{+}{\triangle}} [P] \triangleq [N] \mathbin{\stackrel{+}{\triangle}} [Q]. \qquad (3.3.4)$$

6. Закон вычитания. Если смесь цветов $[M]$ и $[P]$ эквивалентна смеси $[N]$ и $[Q]$ и если цвет $[P]$ эквивалентен цвету $[Q]$, то цвет $[M]$ эквивалентен цвету $[N]$:

$$[M] \mathbin{\stackrel{+}{\triangle}} [P] \triangleq [N] \mathbin{\stackrel{+}{\triangle}} [Q],\ [P] \triangleq [Q] \Rightarrow [M] \triangleq [N]. \qquad (3.3.5)$$

7. Закон транзитивности. Если цвет $[M]$ эквивалентен цвету $[N]$, а цвет $[N]$ — цвету $[P]$, то $[M]$ эквивалентен $[P]$:

$$[M] \triangleq [N],\ [N] \triangleq [P] \Rightarrow [M] \triangleq [P]. \qquad (3.3.6)$$

8. Уравнивание цветов. Справедливо одно из трех соотношений:

а) C единиц цвета $[C]$ уравнивают смесь M единиц $[M]$, N единиц $[N]$ и P единиц $[P]$:

$$C \vartriangle [C] \triangleq M \vartriangle [M] \mathbin{\stackrel{+}{\triangle}} N \vartriangle [N] \mathbin{\stackrel{+}{\triangle}} P \vartriangle [P]; \qquad (3.3.7)$$

б) смесь C единиц $[C]$ и M единиц $[M]$ уравнивает смесь N единиц $[N]$ и P единиц $[P]$:

$$C \vartriangle [C] \mathbin{\stackrel{+}{\triangle}} M \vartriangle [M] \triangleq N \vartriangle [N] \mathbin{\stackrel{+}{\triangle}} P \vartriangle [P]; \qquad (3.3.8)$$

в) смесь C единиц $[C]$, M единиц $[M]$ и N единиц $[N]$ уравнивает P единиц цвета $[P]$:

$$C \vartriangle [C] \mathbin{\stackrel{+}{\triangle}} M \vartriangle [M] \mathbin{\stackrel{+}{\triangle}} N \vartriangle [N] \triangleq P \vartriangle [P]. \qquad (3.3.9$$

3.4. КОЛОРИМЕТРИЯ

Колориметрия — это наука об измерении цветов. Для измерения цветов используются различные системы координат.

Из восьмой аксиомы Грассмана следует, что цвет $[C]$ может быть уравнен смесью трех основных цветов $[P_1]$, $[P_2]$ и $[P_3]$, т. е.

$$[C] \triangleq A_1(C) \vartriangle [P_1] \mathbin{\stackrel{+}{\triangle}} A_2(C) \vartriangle [P_2] \mathbin{\stackrel{+}{\triangle}} A_3(C) \vartriangle [P_3], \qquad (3.4.1)$$

где $A_1(C)$, $A_2(C)$, $A_3(C)$ — уравнивающие величины цвета $[C]$. Так как интенсивности некогерентных источников света складываются линейно, то спектральная плотность смеси равна сумме спектральных плотностей ее компонент. Согласно соотношению (3.4.1), спектральная плотность $C(\lambda)$ может быть заменена экви-

[1]) Вводятся следующие символы: $\triangleq$ — уравнивание (эквивалентность), $\mathbin{\stackrel{+}{\triangle}}$ — аддитивная смесь, $\vartriangle$ — единица смеси.

валентной спектральной плотностью смеси основных цветов, т. е.

$$C(\lambda) \triangleq A_1(C) P_1(\lambda) + A_2(C) P_2(\lambda) + A_3(C) P_3(\lambda) = \sum_{j=1}^{3} A_j(C) P_j(\lambda), \qquad (3.4.2)$$

где $P_j(\lambda)$ — спектральная плотность основного цвета $[P_j]$. Соотношение (3.4.2) имеет простой смысл: спектральные плотности, связанные оператором эквивалентности $\triangleq$, вызывают одинаковое цветовое ощущение. В колориметрии используются координаты цвета, равные по определению

$$T_j(C) = A_j(C)/A_j(W), \qquad (3.4.3)$$

где $A_j(W)$ — уравнивающая величина опорного белого цвета. Подставив выражение (3.4.3) в (3.4.2), получим

$$C(\lambda) \triangleq \sum_{j=1}^{3} T_j(C) A_j(W) P_j(\lambda). \qquad (3.4.4)$$

Из четвертой аксиомы Грассмана следует, что яркость цвета $L(C)$ равна сумме яркостей его основных компонент. Следовательно,

$$L(C) = \int C(\lambda) V(\lambda) d\lambda = \sum_{j=1}^{3} \int A_j(C) P_j(\lambda) V(\lambda) d\lambda, \qquad (3.4.5а)$$

$$L(C) = \sum_{j=1}^{3} \int T_j(C) A_j(W) P_j(\lambda) V(\lambda) d\lambda, \qquad (3.4.5б)$$

где $V(\lambda)$ — относительная световая эффективность. Соотношения (3.4.4) и (3.4.5) являются количественной основой колориметрии.

3.4.1. ПРОВЕРКА МОДЕЛИ ЦВЕТОВОГО ЗРЕНИЯ

Прежде чем продолжить изучение процесса уравнивания цветов, полезно проверить, согласуются ли результаты экспериментов и аксиомы уравнивания цветов с моделью цветового зрения, рассмотренной в разд. 2.5. В этой модели реакции рецепторов трех типов со спектральными чувствительностями $s_1(\lambda)$, $s_2(\lambda)$, $s_3(\lambda)$ представлялись следующими величинами:

$$e_1(C) = \int C(\lambda) s_1(\lambda) d\lambda, \qquad (3.4.6а)$$

$$e_2(C) = \int C(\lambda) s_2(\lambda) d\lambda, \qquad (3.4.6б)$$

$$e_3(C) = \int C(\lambda) s_3(\lambda) d\lambda. \qquad (3.4.6в)$$

Если наблюдатель видит эквивалентную смесь первичных цветов, а не исходный цвет $[C]$, то замена спектральной плотности $C(\lambda)$

эквивалентной спектральной плотностью, определяемой соотношением (3.4.4), не должна приводить к изменению сигналов колбочек $e_i(C)$. Следовательно,

$$e_1(C) = \sum_{j=1}^{3} T_j(C)\, A_j(W) \int P_j(\lambda)\, s_1(\lambda)\, d\lambda, \tag{3.4.7а}$$

$$e_2(C) = \sum_{j=1}^{3} T_j(C)\, A_j(W) \int P_j(\lambda)\, s_2(\lambda)\, d\lambda, \tag{3.4.7б}$$

$$e_3(C) = \sum_{j=1}^{3} T_j(C)\, A_j(W) \int P_j(\lambda)\, s_3(\lambda)\, d\lambda. \tag{3.4.7в}$$

Определив коэффициенты

$$k_{ij} = \int P_j(\lambda)\, s_i(\lambda)\, d\lambda, \tag{3.4.8}$$

можно переписать соотношения (3.4.7) более компактно в матричной форме

$$\begin{bmatrix} e_1(C) \\ e_2(C) \\ e_3(C) \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} A_1(W) & 0 & 0 \\ 0 & A_2(W) & 0 \\ 0 & 0 & A_3(W) \end{bmatrix} \begin{bmatrix} T_1(C) \\ T_2(C) \\ T_3(C) \end{bmatrix} \tag{3.4.9}$$

или еще короче как

$$\mathbf{e}(C) = \mathbf{KAT}(C). \tag{3.4.10}$$

Векторы и матрицы, входящие в это соотношение, определены выражениями (3.4.7)—(3.4.9). Следует отметить, что для заданного набора основных цветов и опорного белого цвета элементы матриц $\mathbf{K}$ и $\mathbf{A}$ оказываются постоянными величинами. Поэтому если известны сигналы колбочек $e_i(C)$ для данного цвета $[C]$, то при условии, что существует матрица $(\mathbf{KA})^{-1}$, соответствующие координаты цвета $T_j(C)$ могут быть вычислены по формуле

$$\mathbf{T}(C) = (\mathbf{KA})^{-1}\, \mathbf{e}(C). \tag{3.4.11}$$

Таким образом, при подходящем выборе величин $T_j(C)$ любой цвет $[C]$ может быть заменен эквивалентной смесью основных цветов; при этом сигналы колбочек останутся неизменными. К сожалению, эти сигналы трудно измерить, поэтому соотношение (3.4.11) нельзя применить для непосредственного вычисления координат цвета. Но это и не было целью приведенного вывода. Наоборот, соотношение (3.4.11) выведено для того, чтобы показать согласованность результатов экспериментов по уравниванию цветов с моделью цветового зрения.

3.4.2. ВЫЧИСЛЕНИЕ КООРДИНАТ ЦВЕТА

Координаты цвета произвольного излучения для заданного набора основных цветов можно вычислить косвенным путем, если для этого набора известны функции сложения, т. е. координаты цвета спектральных (узкополосных) излучений единичной энергии. Типичные функции сложения показаны на рис. 3.4.1. Функции сложения, определяющие колориметрическую систему, обозначаются как $T_{s1}(\lambda)$, $T_{s2}(\lambda)$, $T_{s3}(\lambda)$, где λ — длина волны спектрального цвета.

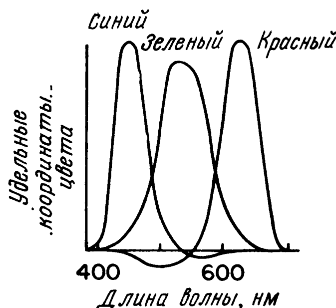


Рис. 3.4.1. Функции сложения для типичных основных цветов — красного, зеленого и синего.

Монохроматический свет $[C_\psi]$ с длиной волны ψ и единичной энергией имеет спектральную плотность $\delta(\lambda - \psi)$. При этом справедливо следующее соотношение:

$$e_i(C_\psi) = \int \delta(\lambda - \psi) s_i(\lambda)\, d\lambda =$$

$$= \sum_{j=1}^{3} A_j(W) P_j(\lambda) T_{sj}(\psi) s_i(\lambda)\, d\lambda. \qquad (3.4.12)$$

Рассмотрим теперь произвольный цвет $[C]$ со спектральной плотностью $C(\lambda)$. $C(\psi)$ единиц этого света при длине волны ψ имеют координаты цвета $T_{s1}(\psi)$, $T_{s2}(\psi)$, $T_{s3}(\psi)$. Следовательно,

$$\int C(\psi) \delta(\lambda - \psi) s_i(\lambda)\, d\lambda =$$

$$= \sum_{j=1}^{3} \int A_j(W) P_j(\lambda) C(\psi) T_{sj}(\psi) s_i(\lambda)\, d\lambda. \qquad (3.4.13)$$

Интегрируя обе части этого равенства по ψ и используя основное свойство дельта-функции, получаем

$$e_i(C) = \iint C(\psi)\,\delta(\lambda - \psi)\,s_i(\lambda)\,d\lambda\,d\psi =$$
$$= \sum_{i=1}^{3} \iint A_j(W)\,P_j(\lambda)\,C(\psi)\,T_{sj}(\psi)\,s_i(\lambda)\,d\psi\,d\lambda. \qquad (3.4.14)$$

Сравнение этого равенства с соотношением (3.4.7) показывает, что координата цвета $[C]$ есть

$$T_j(C) = \int C(\psi)\,T_{sj}(\psi)\,d\psi. \qquad (3.4\ 15)$$

Как видно из рис. 3.4.1, координаты некоторых спектральных цветов могут иметь отрицательные значения. С другой стороны, координата цвета пропорциональна интенсивности основного цвета, которая не может быть отрицательной. Отрицательные значения координат интерпретируются следующим образом: основной цвет, соответствующий отрицательному значению координаты, должен быть прибавлен к исходному цвету, после чего достигается

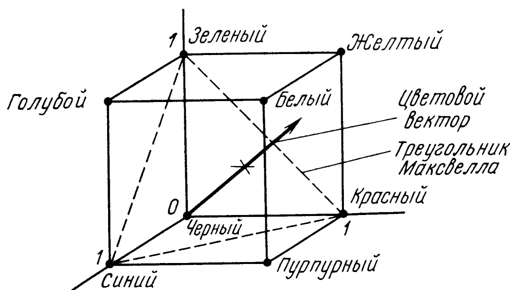


Рис. 3.4.2. Цветовое пространство для типичных основных цветов.

уравнивание смесью двух оставшихся основных цветов. В этом смысле любой цвет может быть уравнен смесью любых основных цветов. Однако цвета с отрицательными координатами не могут быть воспроизведены с помощью реального устройства с заданными основными цветами (например, в приемнике цветного телевидения). К счастью, можно выбрать основные цвета так, чтобы большинство обычных естественных цветов были воспроизводимыми.

Координаты цвета T_1, T_2, T_3 можно рассматривать как координаты точек в некотором трехмерном пространстве (рис. 3.4.2). Данный цвет можно представить также как вектор. На рис. 3.4.2 показан треугольник, называемый треугольником Максвелла. Расстояния от вершин этого треугольника до точки пересечения

цветового вектора с плоскостью треугольника характеризуют цветовой тон и насыщенность цветов.

Часто при описании цвета его яркость не представляет интереса. В таких случаях цветовой тон и насыщенность цвета [*C*]

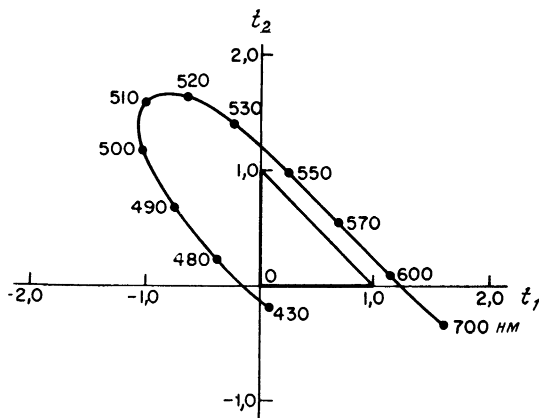


Рис. 3.4.3. График цветностей для типичных основных цветов.

можно выразить через координаты цветности, являющиеся нормированными координатами цвета:

$$t_1 \equiv T_1/(T_1 + T_2 + T_3), \quad (3.4.16а)$$

$$t_2 \equiv T_2/(T_1 + T_2 + T_3), \quad (3.4.16б)$$

$$t_3 \equiv T_3/(T_1 + T_2 + T_3). \quad (3.4.16в)$$

Очевидно, что $t_3 = 1 - t_2 - t_1$ и, следовательно, для описания цветности необходимы только две координаты. На рис. 3.4.3 показан график цветностей для набора типичных основных цветов. С помощью реальных источников света можно физически воспроизводить только те цвета, которые находятся внутри треугольника, определенного тремя основными цветами.

3.4.3. ВЫЧИСЛЕНИЕ ЯРКОСТИ

Координаты цвета определяют данный цвет по отношению к опорному белому цвету. Часто, однако, необходимо знать не относительные, а абсолютные интенсивности основных цветов, воспроизводящих данный цвет. Эти данные можно получить, вычисляя его яркость.

Из выражений (3.4.5) следует, что яркость $L(C)$ цвета [*C*] равна сумме яркостей его основных компонент:

$$L(C) = \sum_{i=1}^{3} T_i(C) \int A_i(W) P_i(\lambda) V(\lambda) \, d\lambda. \quad (3.4.17)$$

Интегралы

$$L(P_i) \equiv \int A_i(W) P_i(\lambda) V(\lambda) d\lambda \qquad (3.4.18)$$

называются яркостными коэффициентами основных цветов. Эти коэффициенты представляют собой яркости основных цветов, уравнивающих опорный белый цвет. Следовательно, яркость цвета C можно записать как

$$L(C) = T_1(C) L(P_1) + T_2(C) L(P_2) + T_3(C) L(P_3). \qquad (3.4.19)$$

Координата цветности цвета $[C]$ определяется как

$$t_1(C) = T_1(C) \Big/ \sum_{j=1}^{3} T_j(C). \qquad (3.4.20)$$

Перемножив соответственно левые и правые части выражений (3.4.19) и (3.4.20), после некоторых преобразований получим

$$T_1(C) = [t_1(C) L(C)] / [t_1(C) L(P_1) + t_2(C) L(P_2) + t_3(C) L(P_3)] \qquad (3.4.21а)$$

и аналогично

$$T_2(C) = [t_2(C) L(C)] / [t_1(C) L(P_1) + t_2(C) L(P_2) + t_3(C) L(P_3)], \qquad (3.4.21б)$$

$$T_3(C) = [t_3(C) L(C)] / [t_1(C) L(P_1) + t_2(C) L(P_2) + t_3(C) L(P_3)]. \qquad (3.4.21в)$$

Таким образом, координаты цвета можно выразить через яркость и координаты цветности.

3.5. ПРЕОБРАЗОВАНИЕ КООРДИНАТ ЦВЕТА

Из выражений (3.4.7) следует, что набор основных цветов можно выбрать многими способами. Если координаты цвета известны для одного набора основных цветов, то простым преобразованием координат можно найти координаты цвета для другого набора [16]. Пусть $[P_1]$, $[P_2]$, $[P_3]$ есть исходный набор основных цветов со спектральными плотностями $P_1(\lambda)$, $P_2(\lambda)$, $P_3(\lambda)$. Интенсивности, уравнивающие опорный белый цвет $[W]$, обозначим через $A_1(W)$, $A_2(W)$ и $A_3(W)$. Рассмотрим теперь другой набор основных цветов $[\tilde{P}_1]$, $[\tilde{P}_2]$, $[\tilde{P}_3]$, имеющих спектральные плотности $\tilde{P}_1(\lambda)$, $\tilde{P}_2(\lambda)$, $\tilde{P}_3(\lambda)$. Опорный белый цвет $[\tilde{W}]$, который может отличаться от исходного опорного белого цвета $[W]$, уравнивается при интенсивностях $\tilde{A}_1(\tilde{W})$, $\tilde{A}_2(\tilde{W})$, $\tilde{A}_3(\tilde{W})$ новых

цветов. Произвольный цвет $[C]$ имеет координаты $T_1(C)$, $T_2(C)$, $T_3(C)$ для исходного набора основных цветов и $\tilde{T}_1(C)$, $\tilde{T}_2(C)$, $\tilde{T}_3(C)$ для нового набора. Из формулы (3.4.10) можно получить матричное соотношение

$$\mathbf{e}(C) = \mathbf{K}\mathbf{A}(W)\mathbf{T}(C) = \tilde{\mathbf{K}}\tilde{\mathbf{A}}(\tilde{W})\tilde{\mathbf{T}}(C). \qquad (3.5.1)$$

Установим теперь единицы измерения новых координат цвета. Взяв вместо цвета $[C]$ новый опорный белый цвет $[\tilde{W}]$, получим

$$\mathbf{e}(\tilde{W}) = \mathbf{K}\mathbf{A}(W)\mathbf{T}\tilde{W} = \tilde{\mathbf{K}}\tilde{\mathbf{A}}(\tilde{W})\tilde{\mathbf{T}}(\tilde{W}), \qquad (3.5.2)$$

где $\tilde{T}_1(\tilde{W}) = \tilde{T}_2(\tilde{W}) = \tilde{T}_3(\tilde{W}) = 1$. Затем подставив в соотношение (3.5.1) новые основные цвета $[P_1]$, $[\tilde{P}_2]$, $[\tilde{P}_3]$, будем иметь

$$\mathbf{e}(\tilde{P}_1) = \mathbf{K}\mathbf{A}(W)\mathbf{T}(\tilde{P}_1) = \tilde{\mathbf{K}}\tilde{\mathbf{A}}(\tilde{W})\tilde{\mathbf{T}}(\tilde{P}_1), \qquad (3.5.3а)$$

$$\mathbf{e}(\tilde{P}_2) = \mathbf{K}\mathbf{A}(W)\mathbf{T}(\tilde{P}_2) = \tilde{\mathbf{K}}\tilde{\mathbf{A}}(\tilde{W})\tilde{\mathbf{T}}(\tilde{P}_2), \qquad (3.5.3б)$$

$$\mathbf{e}(\tilde{P}_3) = \mathbf{K}\mathbf{A}(W)\mathbf{T}(\tilde{P}_3) = \tilde{\mathbf{K}}\tilde{\mathbf{A}}(\tilde{W})\tilde{\mathbf{T}}(\tilde{P}_3), \qquad (3.5.3в)$$

где

$$\tilde{T}(\tilde{P}_1) = \begin{bmatrix} 1/\tilde{A}_1(\tilde{W}) \\ 0 \\ 0 \end{bmatrix}, \quad \tilde{T}(\tilde{P}_2) = \begin{bmatrix} 0 \\ 1/\tilde{A}_2(\tilde{W}) \\ 0 \end{bmatrix},$$

$$\tilde{T}(\tilde{P}_3) = \begin{bmatrix} 0 \\ 0 \\ 1/\tilde{A}_3(W) \end{bmatrix}.$$

Совместное решение системы матричных уравнений (3.5.1), (3.5.2) и (3.5.3) дает требуемое соотношение между координатами цвета для исходного и нового наборов основных цветов:

$$\begin{bmatrix} \tilde{T}_1(C) \\ \tilde{T}_2(C) \\ \tilde{T}_3(C) \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} T_1(C) & T_1(\tilde{P}_2) & T_1(\tilde{P}_3) \\ T_2(C) & T_2(\tilde{P}_2) & T_2(\tilde{P}_3) \\ T_3(C) & T_3(\tilde{P}_2) & T_3(\tilde{P}_3) \end{vmatrix} \Big/ \begin{vmatrix} T_1(\tilde{W}) & T_1(\tilde{P}_2) & T_1(\tilde{P}_3) \\ T_2(\tilde{W}) & T_2(\tilde{P}_2) & T_2(\tilde{P}_3) \\ T_3(\tilde{W}) & T_3(\tilde{P}_2) & T_3(\tilde{P}_3) \end{vmatrix} \\ \\ \begin{vmatrix} T_1(\tilde{P}_1) & T_1(C) & T_1(\tilde{P}_3) \\ T_2(\tilde{P}_1) & T_2(C) & T_2(\tilde{P}_3) \\ T_3(\tilde{P}_1) & T_3(C) & T_3(\tilde{P}_3) \end{vmatrix} \Big/ \begin{vmatrix} T_1(\tilde{P}_1) & T_1(\tilde{W}) & T_1(\tilde{P}_3) \\ T_2(\tilde{P}_1) & T_2(\tilde{W}) & T_2(\tilde{P}_3) \\ T_3(\tilde{P}_1) & T_3(\tilde{W}) & T_3(\tilde{P}_3) \end{vmatrix} \\ \\ \begin{vmatrix} T_1(\tilde{P}_1) & T_1(\tilde{P}_2) & T_1(C) \\ T_2(\tilde{P}_1) & T_2(\tilde{P}_2) & T_2(C) \\ T_3(\tilde{P}_1) & T_3(\tilde{P}_2) & T_3(C) \end{vmatrix} \Big/ \begin{vmatrix} T_1(\tilde{P}_1) & T_1(\tilde{P}_2) & T_1(\tilde{W}) \\ T_2(\tilde{P}_1) & T_2(\tilde{P}_2) & T_2(\tilde{W}) \\ T_3(\tilde{P}_1) & T_3(\tilde{P}_2) & T_3(\tilde{W}) \end{vmatrix} \end{bmatrix} \cdot \qquad (3.5.4)$$

Можно переписать соотношение (3.5.4), используя координаты цветности $t_i(\tilde{P}_1)$, $t_i(\tilde{P}_2)$, $t_i(\tilde{P}_3)$ новых основных цветов в координатной системе исходных основных цветов. В этом случае

$$\begin{bmatrix} \tilde{T}_1(C) \\ \tilde{T}_2(C) \\ \tilde{T}_3(C) \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} T_1(C) \\ T_2(C) \\ T_3(C) \end{bmatrix}, \qquad (3.5.5)$$

где

$$m_{ij} = \Delta_{ij}/\Delta_i,$$

$$\Delta_1 = T_1(\tilde{W})\,\Delta_{11} + T_2(\tilde{W})\,\Delta_{12} + T_3(\tilde{W})\,\Delta_{13},$$

$$\Delta_2 = T_1(\tilde{W})\,\Delta_{21} + T_2(\tilde{W})\,\Delta_{22} + T_3(\tilde{W})\,\Delta_{23},$$

$$\Delta_3 = T_1(\tilde{W})\,\Delta_{31} + T_2(\tilde{W})\,\Delta_{32} + T_3(\tilde{W})\,\Delta_{33},$$

$$\Delta_{11} = t_2(\tilde{P}_2)\,t_3(\tilde{P}_3) - t_3(\tilde{P}_2)\,t_2(\tilde{P}_3),$$

$$\Delta_{12} = t_3(\tilde{P}_2)\,t_1(\tilde{P}_3) - t_1(\tilde{P}_2)\,t_3(\tilde{P}_3),$$

$$\Delta_{13} = t_1(\tilde{P}_2)\,t_2(\tilde{P}_3) - t_2(\tilde{P}_2)\,t_1(\tilde{P}_3),$$

$$\Delta_{21} = t_3(\tilde{P}_1)\,t_2(\tilde{P}_3) - t_2(\tilde{P}_1)\,t_3(\tilde{P}_3),$$

$$\Delta_{22} = t_1(\tilde{P}_1)\,t_3(\tilde{P}_3) - t_3(\tilde{P}_1)\,t_1(\tilde{P}_3),$$

$$\Delta_{23} = t_2(\tilde{P}_1)\,t_1(\tilde{P}_3) - t_1(\tilde{P}_1)\,t_2(\tilde{P}_3),$$

$$\Delta_{31} = t_2(\tilde{P}_1)\,t_3(P_2) - t_3(\tilde{P}_1)\,t_2(\tilde{P}_2),$$

$$\Delta_{32} = t_3(\tilde{P}_1)\,t_1(\tilde{P}_2) - t_1(\tilde{P}_1)\,t_3(\tilde{P}_2),$$

$$\Delta_{33} = t_1(\tilde{P}_1)\,t_2(\tilde{P}_2) - t_2(\tilde{P}_1)\,t_1(\tilde{P}_2).$$

3.6. СИСТЕМЫ КООРДИНАТ ЦВЕТА

Было показано, что цвет $[C]$ может быть описан координатами цвета $T_1(C)$, $T_2(C)$, $T_3(C)$ для данного набора основных цветов. С другой стороны, цвет можно определить координатами цветности $t_1(C)$, $t_2(C)$ и яркостью $L(C)$. Можно также описать цвет, используя какую-либо линейную или нелинейную обратимую функцию координат цвета или координат цветности и яркости. Из выражения (3.5.5) видно, что линейное преобразование координат цвета есть просто переход к новому набору основных цветов. В приложении 2 приведены формулы преобразования координат цвета и цветности для различных систем координат.

Для количественного описания цветов предложено много разных систем координат. Ниже рассматриваются те из них, которые представляют исторический и теоретический интерес.

Система координат спектральных основных цветов МКО

В 1931 г. МКО разработала стандартный набор монохроматических основных цветов· красный с длиной волны 700 нм, зеленый — 546,1 нм и синий — 435,8 нм [4]. Единицы измерения координат цвета выбраны так, чтобы координаты R_c, G_c, B_c белого света с равномерной спектральной плотностью в видимой части спектра были одинаковыми. Набор основных цветов определяется

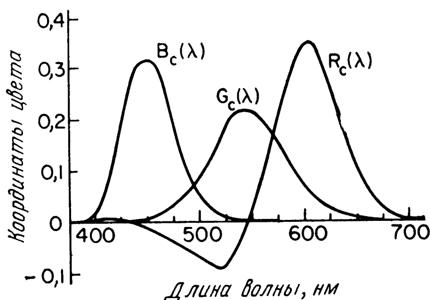


Рис. 3.6.1. Функции сложения системы координат спектральных основных цветов МКО (красный — 700 нм, зеленый — 546,1 нм, синий — 435,8 нм) [11].

кривыми сложения для спектральных цветов, приведенными на рис. 3.6.1. Эти кривые получены в экспериментах по уравниванию цветов с большим числом наблюдателей. По результатам экспериментов был определен так называемый стандартный наблюдатель МКО. Данные для поля зрения 2° были опубликованы в 1931 г. Затем были получены результаты для поля размером 10°. В телевидении и фототелеграфии лучше использовать данные для поля 2°. На рис. 3.6.2 представлен график цветностей в системе координат спектральных основных цветов МКО, а также цвета люминофоров телевизионного приемника принятой в США системы цветного телевидения НТСЦ. Треугольник, определяемый цветами люминофоров, охватывает цветности всех воспроизводимых цветов.

Система координат приемника НТСЦ

В телевизионных приемниках США используются кинескопы с тремя люминофорами — красным, зеленым и синим [14]. Система координат приемника НТСЦ, определяемая цветами люминофоров, может быть связана с системой координат спектральных основных цветов МКО простым линейным преобразованием. На рис. 3.6.3 приведен график цветностей в системе координат приемника НТСЦ. В этой системе единицы измерения координат цвета нормированы так, что значения координат, при которых уравни-

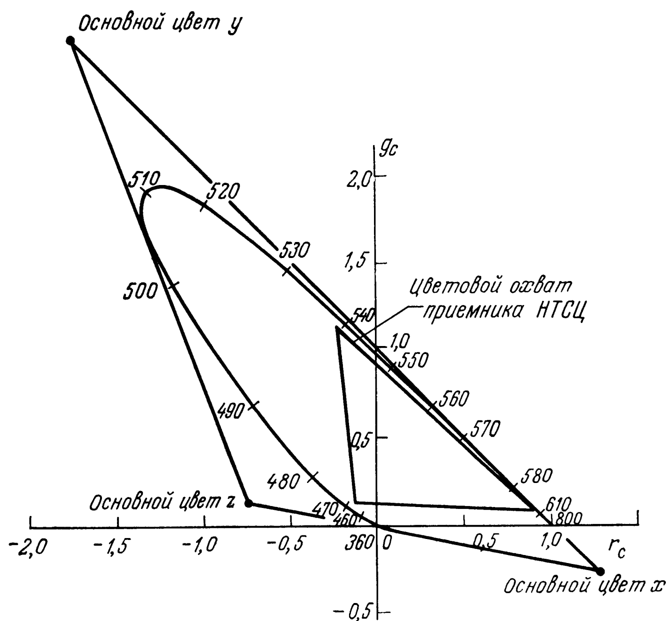


Рис. 3.6.2. График цветностей в системе координат спектральных основных цветов МКО [11].

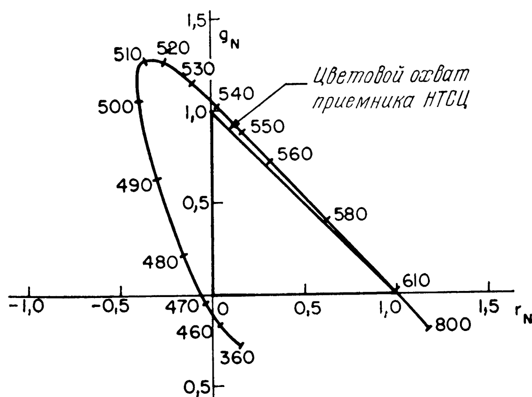


Рис. 3.6.3. График цветностей в системе координат приемника НТСЦ.

вается опорный белый цвет, одинаковы. Люминофоры приемника НТСЦ не являются источниками монохроматического света, поэтому определяемый ими цветовой охват (совокупность воспроизводимых цветов) уже, чем при использовании спектральных основных цветов МКО.

Система координат XYZ МКО

Система координат спектральных основных цветов МКО имеет один недостаток для колориметрических расчетов: координаты цвета иногда оказываются отрицательными. Столкнувшись с этой трудностью, МКО разработала систему координат с искусственными основными цветами, в которой координаты цвета спектральных цветов являются положительными [4]. Искусственные основ-

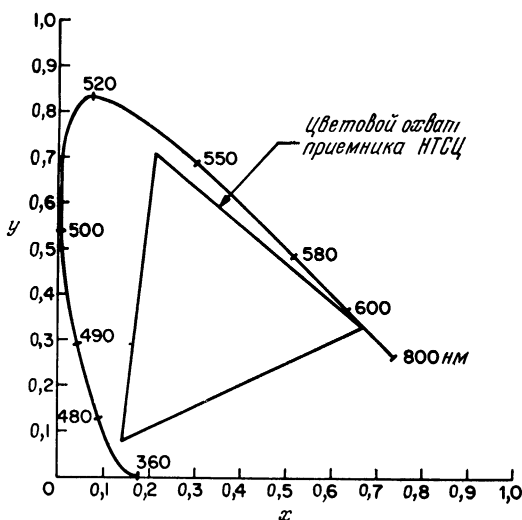


Рис. 3.6.4. График цветностей в системе координат *XYZ* МОК [4].

ные цвета показаны на рис. 3.6.2. Они выбираются так, чтобы координата *Y* была эквивалентна яркости *L* (*C*) цвета *C*. На рис. 3.6.4 представлен график цветностей в системе *XYZ* МКО при опорном белом свете с равномерной спектральной плотностью.

Система координат передаваемых сигналов НТСЦ

В системе цветного телевидения НТСЦ, разработанной в США, передаются три координаты цвета *Y*, *I*, *Q* [14]. Координата *Y* совпадает с координатой *Y* системы *XYZ*; она соответствует ярко-

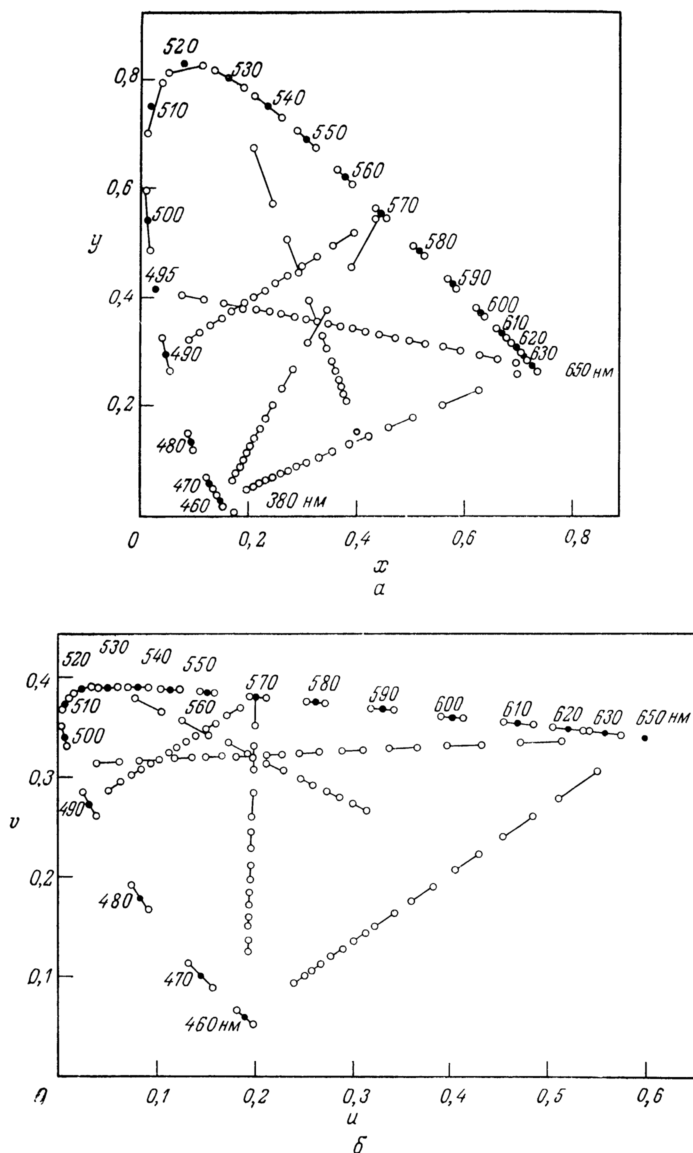


Рис. 3.6.5. Сравнение едва заметных цветовых разностей в системах координат *XYZ* и *UVW*. Величины разностей увеличены в 10 раз [12].

а — цветовые разности на графике цветностей (x, y); *б* — цветовые разности на графике цветностей (u, v).

сти. Остальные две координаты I и Q вместе описывают цветовой тон и насыщенность. Причинами передачи координат Y, I, Q вместо координат R_N, G_N, B_N непосредственно с выхода передающей камеры являются следующие: 1) сигнал Y может быть использован существующими телевизионными приемниками одноцветного изображения и 2) полосу частот сигналов I и Q можно сократить без заметных искажений изображений. Применив такое сокращение и остроумный способ модуляции, удалось передавать полный аналоговый сигнал цветного телевидения в той же полосе частот, что и при одноцветном изображении.

Равноконтрастная система координат МКО

Желательно иметь такую систему координат, чтобы равным изменениям координат цветности соответствовали равные изменения в ощущении цвета. На графике цветностей (рис. 3.6.5, *а*) показаны разности цветов, которые воспринимаются одинаково [12, 17]. Этот график, а также другие экспериментальные резуль-

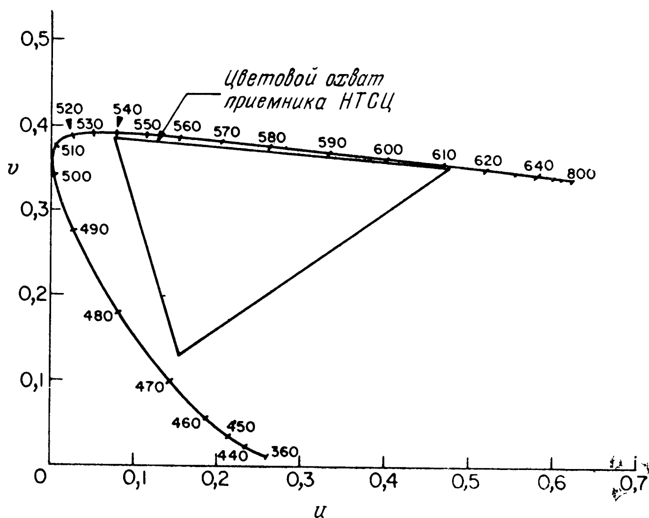


Рис. 3.6.6. Равноконтрастный график цветностей [11].

таты свидетельствуют о том, что человеческий глаз наиболее чувствителен к изменению синего цвета, умеренно чувствителен к изменению красного и обладает наименьшей чувствительностью к изменениям зеленого цвета. В 1960 г. МКО приняла равноконтрастную систему координат, в которой с хорошим приближением равные изменения координат цветности соответствуют едва заметным изменениям цветового тона и насыщенности. На рис. 3.6.5, *б* приведены данные рис. 3.6.5, *а* в равноконтрастных координатах.

Переход от координат X, Y, Z к равноконтрастным координатам осуществляется линейным преобразованием. Координаты цветности в обеих системах связаны следующими соотношениями [18]:

$$u = 4x/(-2x + 12y + 3), \quad (3.6.1а)$$

$$v = 6y/(-2x + 12y + 3). \quad (3.6.1б)$$

График цветностей в равноконтрастной системе координат приведен на рис. 3.6.6.

Система координат U*V*W*

Система координат $U^*V^*W^*$ есть развитие системы координат UVW с целью получения цветового пространства, в котором единичные изменения цветности и яркости воспринимаются одинаково. Координаты U^*, V^*, W^* по определению [19] равны

$$U^* = 13W^* (u - u_0), \quad (3.6.2а)$$

$$V^* = 13W^* (v - v_0), \quad (3.6.2б)$$

$$W^* = 25\,(100Y)^{1/3} - 17, \quad (3.6.2в)$$

причем единицы измерения яркости Y выбраны так, что яркость меняется от 0 до 1, а u_0 и v_0 — координаты цветности опорного белого цвета.

Система координат SθW*

Координаты S, θ, W^* есть просто полярные координаты для системы $U^*V^*W^*$. По определению они равны [11]

$$S = [(U^*)^2 + (V^*)^2]^{1/2} = 13W^* [(u - u_0)^2 + (v - v_0)^2]^{1/2}, \quad (3.6.3а)$$

$$\theta = \operatorname{arctg} [V^*/U^*] = \operatorname{arctg} [(v - v_0)/(u - u_0)]. \quad (3.6.3б)$$

Координата S определяет насыщенность цвета, а θ — цветовой тон.

Система координат Lab

Система координат Lab, будучи еще достаточно простой для колориметрических расчетов, обеспечивает относительно точное представление цветов в соответствии с системой цветов Мюнселла [20]. Координаты цвета в этой системе равны

$$L = 25\,(100Y/Y_0)^{1/3} - 16, \quad (3.6.4а)$$

$$a = 500\,[(X/X_0)^{1/3} - (Y/Y_0)^{1/3}], \quad (3.6.4б)$$

$$b = 200\,[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3}], \quad (3.6.4в)$$

где X_0, Y_0, Z_0 — координаты опорного белого цвета в системе XYZ. Координата L определяет яркость цвета, a — соотношение

красного и зеленого цветов, b — соотношение синего и желтого. Многие колориметры, выпускаемые промышленностью, дают значения этих координат.

Система координат Карунена—Лоэва

Систему координат спектральных основных цветов, системы XYZ, UVW и другие можно рассматривать как результат линейного преобразования системы координат приемника НТСЦ. Координаты цвета R_N, G_N, B_N приемника НТСЦ оказываются сильно коррелированными друг с другом [21]. При разработке эффективных методов квантования и кодирования цветных изображений удобно иметь дело с некоррелированными компонентами. Если известна ковариационная матрица величин R_N, G_N, B_N, то можно построить систему ортогональных некоррелированных координат, используя преобразование Карунена—Лоэва. Матрица преобразования $M = [m_{ij}]$ состоит из собственных векторов ковариационной матрицы и определяется следующим соотношением:

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{12} & u_{22} & u_{23} \\ u_{13} & u_{23} & u_{33} \end{bmatrix} \begin{bmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} = \begin{bmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{bmatrix}, \tag{3.6.5}$$

где Λ_1, Λ_2, Λ_3 — собственные значения ковариационной матрицы и

$$u_{11} \equiv E\{(R_N - \overline{R}_N)^2\}, \tag{3.6.6а}$$

$$u_{22} \equiv E\{(G_N - \overline{G}_N)^2\}, \tag{3.6.6б}$$

$$u_{33} \equiv E\{(B_N - \overline{B}_N)^2\}, \tag{3.6.6в}$$

$$u_{12} \equiv E\{(R_N - \overline{R}_N)(G_N - \overline{G}_N)\}, \tag{3.6.6г}$$

$$u_{13} \equiv E\{(R_N - \overline{R}_N)(B_N - \overline{B}_N)\}, \tag{3.6.6д}$$

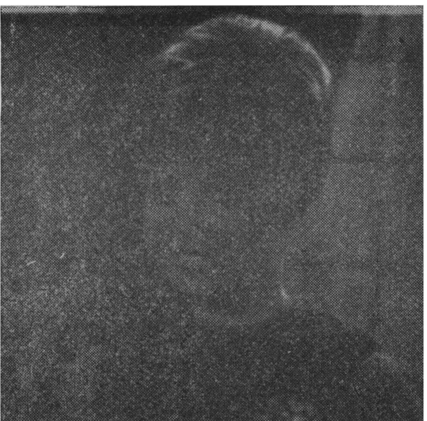
$$u_{23} \equiv E\{(G_N - \overline{G}_N)(B_N - \overline{B}_N)\}. \tag{3.6.6е}$$

Система координат колбочек сетчатки

Как отмечалось в гл. 2, при рассмотрении моделей цветового зрения человека были проведены косвенные измерения спектральных чувствительностей $s_1(\lambda)$, $s_2(\lambda)$, $s_3(\lambda)$ колбочек трех типов.

a
R_N

G_N


B_N

б
G_1

G_2

G_3

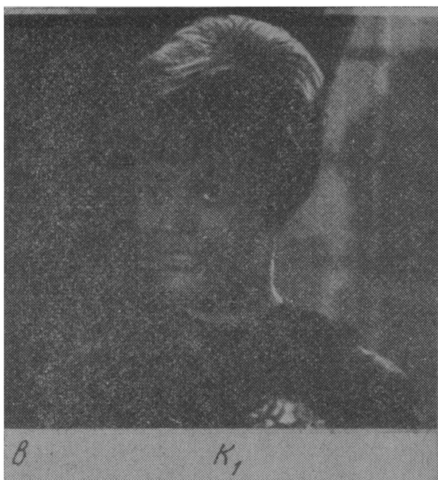
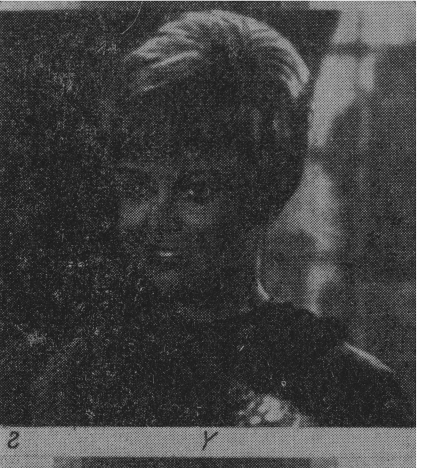


Рис. 3.6.7. Компоненты цветного изображения.

а — система координат приемника НТСЦ; *б* — система координат колбочек; *в* — система координат Карунена—Лоэва:



г Y

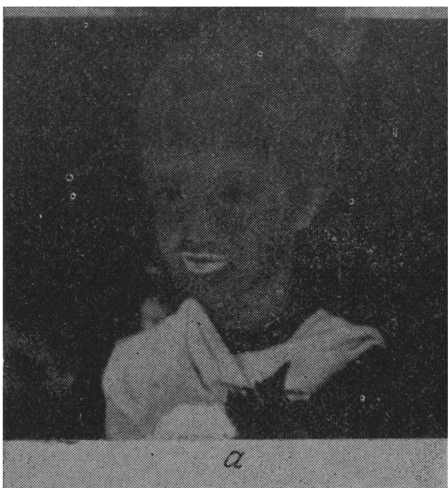
I

Q

д U

V

W



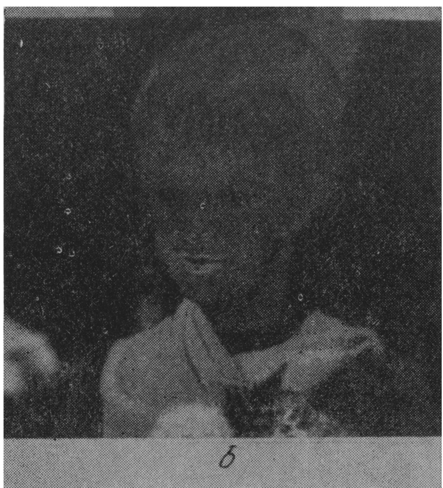


Рис. 3.6.7. *(продолжение)*

г — система координат передаваемых сигналов **НТСЦ**; *д* — равноконтрастная система координат: *е* — система координат *Lab*.

Оказалось, что спектральные чувствительности связаны линейно с функциями сложения, полученными по данным колориметрических экспериментов. Следовательно, сигналы колбочек T_1, T_2, T_3 можно рассматривать как координаты цвета. Эти координаты связаны с координатами X, Y, Z следующим линейным преобразованием [22]:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} 0 & 1,0 & 0 \\ -0,46 & 1,359 & 0,101 \\ 0 & 0 & 1,0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \qquad (3.6.7)$$

На рис. 3.6.7 приведены компоненты цветного изображения для нескольких координатных систем. Следует отметить, что красная, зеленая и синяя компоненты сильно коррелированы. В некоторых координатных системах одна из компонент содержит большую часть энергии изображения, а остальные кажутся менее детальными.

3.7. ЦВЕТОВОЕ ТЕЛО

Как было показано в предыдущем разделе, существует много систем координат, которые можно использовать для описания цветов. К сожалению, нет способа определить, какая из этих систем является оптимальной для большинства приложений. Для выбора наилучшей системы следует сформулировать предъявляемые к ней требования, а затем проверять, в какой степени им удовлетворяют те или иные системы. При количественном сравнении координатных систем следует учитывать ограничения, накладываемые на способы представления цветов, и методы обнаружения различий между цветами.

Для любой системы воспроизведения естественных цветов существуют два физических ограничения. Во-первых, интенсивность света, излучаемого любыми источниками, может быть только положительной. Следовательно, если некоторый цвет имеет отрицательные координаты, он не может быть воспроизведен в виде чисто аддитивной смеси основных цветов. Второе ограничение связано с диапазоном изменения яркости света, прошедшего через предмет или отраженного от него.

3.7.1. ОГРАНИЧЕНИЕ ПОЛЯРНОСТИ КООРДИНАТ ЦВЕТА

Рассмотрим датчик сигналов цветного изображения, на выходе которого получаются координаты $R_N(C)$, $G_N(C)$, $B_N(C)$ в системе основных цветов $[R_N]$, $[G_N]$, $[B_N]$ телевизионного приемника НТСЦ. Этот приемник может воспроизвести любой цвет, имеющий

неотрицательные координаты. Такие цвета лежат внутри куба цветового охвата, показанного на рис. 3.7.1, *а*. Вершины куба соответствуют главным цветам приемника НТСЦ — красному, зеленому, синему, голубому, пурпурному, желтому, черному и

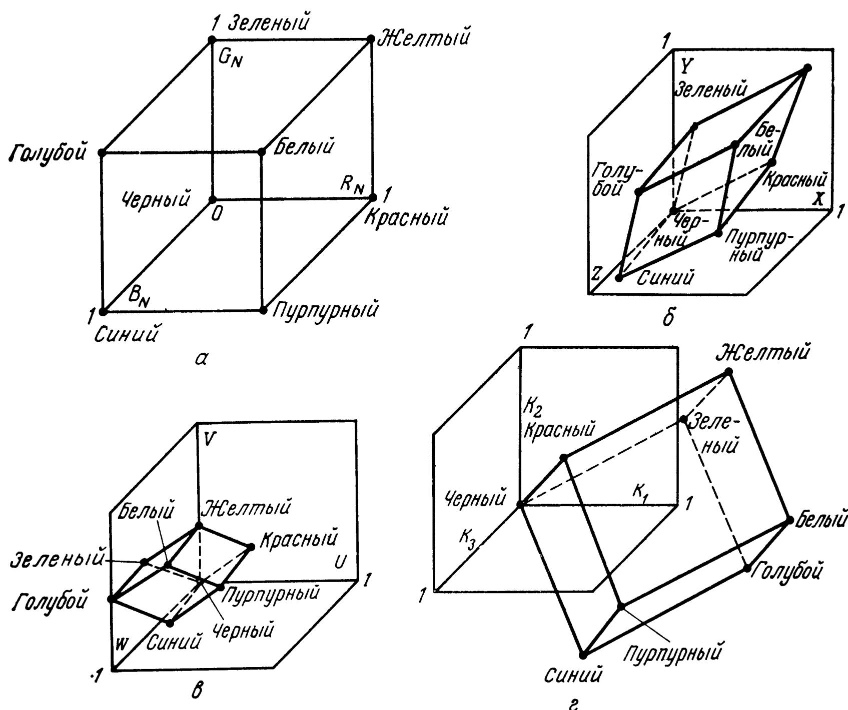


Рис. 3.7.1. Цветовой охват приемника НТСЦ.

а — система координат R_N, G_N, B_N; *б* — система координат XYZ; *в* — система координат UVW; *г* — система координат Карунена—Лоэва.

белому. Посредством линейного преобразования перейдем к другой системе координат цвета T_1 (C), T_2 (C), T_3 (C):

$$\begin{bmatrix} T_1(C) \\ T_2(C) \\ T_3(C) \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} R_N(C) \\ G_N(C) \\ B_N(C) \end{bmatrix}. \qquad (3.7.1)$$

В табл. 3.7.1 приведены координаты цвета и координаты цветности главных цветов приемника НТСЦ для нескольких координатных систем. Цветовой охват, который обеспечивается люминофорами приемника НТСЦ, показан на рис. 3.7.1 в различных системах

Таблица 3.7.1

Координаты цвета и координаты цветности главных цветов приемника НТСЦ для нескольких координатных систем

	Красный	Зеленый	Синий	Голубой	Пурпурный	Желтый	Белый	Черный
R_N	1,000	0,000	0,000	0,000	1,000	1,000	1,000	0,000
G_N	0,000	1,000	0,000	1,000	0,000	1,000	1,000	0,000
B_N	0,000	0,000	1,000	1,000	1,000	0,000	1,000	0,000
r_N	1,000	0,000	0,000	0,000	0,500	0,500	0,333	0,333
g_N	0,000	1,000	0,000	0,500	0,000	0,500	0,333	0,333
b_N	0,000	0,000	1,000	0,500	0,500	0,000	0,333	0,333
X	0,607	0,174	0,201	0,375	0,808	0,781	0,982	0,000
Y	0,299	0,587	0,114	0,701	0,413	0,886	1,000	0,000
Z	0,000	0,066	1,117	1,183	1,117	0,066	1,183	0,000
x	0,670	0,210	0,140	0,166	0,346	0,451	0,310	0,310
y	0,330	0,710	0,080	0,310	0,177	0,511	0,316	0,316
z	0,000	0,080	0,780	0,524	0,477	0,038	0,374	0,374
R_c	0,207	−0,026	−0,026	−0,052	0,180	0,181	0,154	0,000
G_c	0,020	0,133	0,028	0,161	0,048	0,153	0,181	0,000
B_c	0,000	0,010	0,200	0,210	0,199	0,010	0,210	0,000
r_c	0,912	−0,219	−0,132	−0,164	0,421	0,525	0,283	0,283
g_c	0,089	1,130	0,139	0,506	0,113	0,446	0,333	0,333
b_c	0,000	0,089	0,993	0,658	0,466	0,029	0,384	0,384
U	0,405	0,116	0,134	0,250	0,539	0,521	0,655	0,000
V	0,299	0,587	0,114	0,701	0,413	0,886	1,000	0,000
W	0,145	0,827	0,629	1,456	0,774	0,972	1,601	0,000
u	0,477	0,076	0,153	0,104	0,312	0,219	0,201	0,201
v	0,352	0,384	0,130	0,291	0,239	0,373	0,307	0,307
w	0,171	0.540	0,717	0,605	0,449	0,408	0,492	0,492
Y	0,299	0,587	0,114	0,701	0,413	0,886	1,000	0,000
I	0,596	−0,273	−0,322	−0,595	0,272	0,322	0,000	0,000
Q	0,212	−0,522	0,315	−0,207	0,526	−0,312	0,000	0,000
U^*	217,358	−130,319	−24,558	−108,566	100,385	22,144	0,000	0,000
V^*	35,461	79,703	−90,508	−17,913	−61,336	80,125	0,000	0,000
W^*	60,594	80,160	39,265	86,081	69,415	94,451	99,040	−17,000
S	220,232	152,760	93,781	110,034	117,640	83,128	0,000	0,000
θ	0,162	2,609	4,445	3,304	5,733	1,280	0,000	0,000
W^*	60,594	80,160	39,265	86,081	69,415	94,451	99,040	−17,000
L	61,594	81,160	40,265	87,081	70,415	95,451	100,040	−16,000
	91,577	−137,815	52,227	−81,411	96,177	−16,977	0,000	0,000
b	133,738	91,038	−99,233	−22,335	−47,269	115,670	0,000	0,000

координат. Отметим, что в координатной системе Карунена—Лоэва (рис. 3.7.1, *г*) цветовой охват представляется кубом, повернутым относительно осей координат.

Данный цвет может быть описан значениями цветности

$$t_1(C) = T_1(C)/[T_1(C) + T_2(C) + T_3(C)], \qquad (3.7.2\text{а})$$

$$t_2(C) = T_2(C)/[T_1(C) + T_2(C) + T_3(C)] \qquad (3.7.2\text{б})$$

и яркости

$$Y(C) = L(C) = n_{21}R_N(C) + n_{22}G_N(C) + n_{23}B_N(C), \qquad (3.7.2\text{в})$$

где координаты цвета $T_i(C)$ задаются соотношением (3.7.1), а постоянные n_{ij} определены в приложении 2. Преобразование (3.7.2)

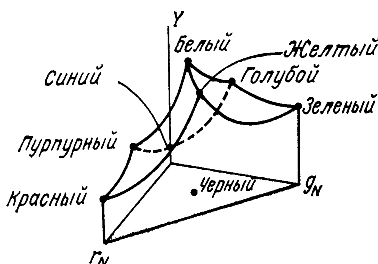


Рис. 3.7.2. Цветовой охват приемника НТСЦ в координатах цветности и яркости (r_N, g_N, Y).

координат цвета в координаты цветности является нелинейным. На рис. 3.7.2 показан цветовой охват приемника НТСЦ в координатах r_N, g_N, Y.

3.7.2. ОГРАНИЧЕНИЕ ДИАПАЗОНА ЯРКОСТЕЙ

Большинство естественных изображений создается светом, отраженным от непрозрачных предметов или прошедшим через прозрачные среды. Диапазон изменения яркости такого несамосветящегося предмета ограничен и зависит от его цвета. Мак-Адам [23, 24] доказал, что цветовое тело максимального объема реализуется в том случае, когда предметы имеют коэффициенты отражения или пропускания, представленные на рис. 3.7.3. Такие предметы физически не существуют; следовательно, цветовое тело реальных предметов имеет меньший объем. Поясним это, ограничившись для простоты только случаем предметов, отражающих свет. Эти же рассуждения останутся справедливыми и в случае прозрачных предметов.

Рассмотрим цвет $[C]$, заданный координатами цвета

$$X = \int r_0(\lambda) E(\lambda) X_s(\lambda)\, d\lambda, \qquad (3.7.3а)$$

$$Y = \int r_0(\lambda) E(\lambda) Y_s(\lambda)\, d\lambda, \qquad (3.7.3б)$$

$$Z = \int r_0(\lambda) E(\lambda) Z_s(\lambda)\, d\lambda, \qquad (3.7.3в)$$

где $r_0(\lambda)$ — спектральный коэффициент отражения, $E(\lambda)$ — спектральная плотность излучения света, падающего на предмет,

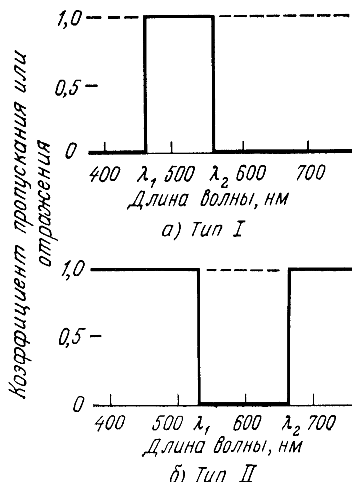


Рис. 3.7.3. Оптимальные спектральные коэффициенты отражения или пропускания несамосветящихся предметов.

$X_s(\lambda)$, $Y_s(\lambda)$, $Z_s(\lambda)$ — функции сложения. Визуальная эффективность β определяется как отношение яркости предмета с коэффициентом отражения $r_0(\lambda)$ к яркости предмета с единичным коэффициентом отражения, т. е.

$$\beta \equiv \int r_0(\lambda) E(\lambda) Y_s(\lambda)\, d\lambda \Big/ \int E(\lambda) Y_s(\lambda)\, d\lambda. \qquad (3.7.4)$$

Если предмет имеет коэффициент отражения такого типа, как показано на рис. 3.7.3, *а*, его яркость равна

$$Y = \int_{\lambda_1}^{\lambda_2} E(\psi_k) Y_s(\psi_k)\, d\psi_k = \beta \int_0^{\infty} E(\psi_k) Y_s(\psi_k)\, d\psi_k. \qquad (3.7.5)$$

Предположим, что известна спектральная плотность $E(\lambda)$ энергии падающего света. Тогда можно вычислить яркость предмета, если известно значение визуальной эффективности β. Для задан-

ного минимального значения длины волны λ_1 максимальное значение λ_2 можно найти численным методом из уравнения

$$\int_0^{\lambda_2} E(\psi_k) Y_s(\psi_k)\, d\psi_k = Y + \int_0^{\lambda_1} E(\psi_k) Y_s(\psi_k)\, d\psi_k. \qquad (3.7.6)$$

Используя найденное значение λ_2, можно вычислить координаты цвета X, Z:

$$X = \int_{\lambda_1}^{\lambda_2} E(\lambda_k) X_s(\lambda_k)\, d\lambda_k, \qquad (3.7.7)$$

$$Z = \int_{\lambda_1}^{\lambda_2} E(\lambda_k) Z_s(\lambda_k)\, d\lambda_k. \qquad (3.7.8)$$

Координаты цветности x и y могут быть найдены по координатам цвета. Если эту процедуру повторить для всевозможных значений

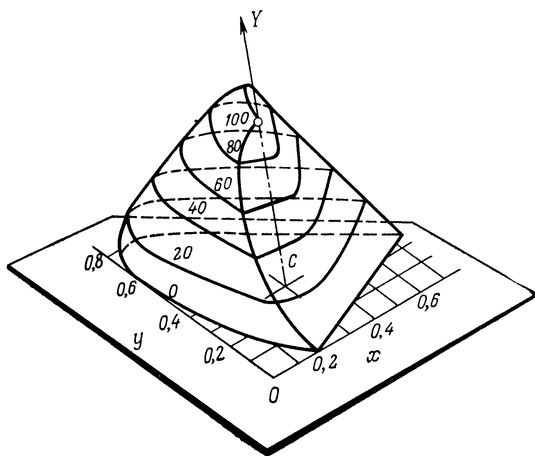


Рис. 3.7.4. Цветовое тело предметов в координатах цветности и яркости (X, Y, Z) [23].

λ_1 при заданных значениях визуальной эффективности β, то можно получить поверхность, ограничивающую тело возможных цветов предметов, в координатах цвета или координатах цветности и яркости. На рис. 3.7.4 показана такая поверхность в координатах цветности и яркости при освещении предметов источником белого света. Отметим, что максимум яркости соответствует белой поверхности.

ЛИТЕРАТУРА

1. Merrit T. P., Hall F. F., Jr., Blackbody Radiation, *Proc. IRE*, **47**, 9, 1435—1442 (September 1959).
2. Malitson H. H., The Solar Energy Spectrum, *Sky and Telescope*, **29**, 4, 162—165 (March 1965).
3. Larabee R. D., Spectral Emissivity of Tungsten, *J. Opt. Soc. Am.*, **49**, 6, 619—625 (June 1959).
4. The Science of Color, Crowell, New York, 1953.
5. Fink D. G., Ed., Television Engineering Handbook, McGraw-Hill, New York, 1957.
6. Walsh J. W. T., Photometry, Constable, London, 1953.
7. Morgan J., Introduction to Geometrical and Physical Optics, McGraw-Hill, New York, 1953.
8. Born M., Wolf E., Principles of Optics, Pergamon, New York, 1970. [Имеется перевод: Борн М., Вольф Э., Основы оптики. — М.: Наука, 1973.]
9. Weaver K. S., The Visibility of Radiation at Low Intensities, *J. Opt. Soc. Am.*, **27**, 1, 39—43 (January 1937).
10. Wyszecki G., Stiles W. S., Color Science, Wiley, New York, 1957.
11. Hunt R. W. G., The Reproduction of Color, Wiley, New York, 1957.
12. Wright W. D., The Measurement of Color, Adam Hilger, London, 1944, pp. 204—205.
13. Enyord R. A., Ed., Color: Theory and Imaging Systems, Society of Photographic Scientists and Engineers, Washington D. C., 1973.
14. Bingley F. J., Color Vision and Colorimetry, Television Engineering Handbook, Fink D. G., Ed., McGraw-Hill, New York, 1957.
15. Grassman H., On the Theory of Compound Colours, *Phil. Mag.*, Ser. 4,**7**, 254—264 (April 1854).
16. Wintringham W. T., Color Television and Colorimetry, *Proc. IRE*, **39**, 10, 1135—1172 (October 1951).
17. MacAdam D. L., Visual Sensitivities to Small Color Differences in Daylight, *J. Opt. Soc. Am.*, **32**, 5, 247—274 (May 1942).
18. MacAdam D. L., Projective Transformations of ICI Color Specifications, *J. Opt. Soc. Am.*, **27**, 8, 294—299 (August 1937).
19. Wyszecki G., Proposal for a New Color-Difference Formula, *J. Opt. Soc. Am.*, **53**, 11, 1318—1319 (November 1963).
20. C. I. E. Colorimetry Committee Proposal for Study of Color Spaces, Tech. Note, *J. Opt. Soc. Am.*, **64**, 6, 896—897 (June 1974).
21. Pratt W. K., Spatial Transform Coding of Color Images, *IEEE Trans. Commun. Tech.*, COM-**19**, 12, 980—992 (December 1971).
22. Judd D. B., Standard Response Functions for Protanopic and Deuteranopic Vision, *J. Opt. Soc. Am.*, **35**, 3, 199—221 (March 1945).
23. MacAdam D. L., The Theory of the Maximum Visual Efficiency of Colored Materials, *J. Opt. Soc. Am.*, **25**, 8, 249—252 (August 1935).
24. MacAdam D. L., Maximum Visual Efficiency of Colored Materials, *J. Opt. Soc. Am.*, **25**, 11, 361—367 (November 1935).

Часть 2

Представление изображений в цифровой форме

При цифровой обработке изображений производится преобразование непрерывного (аналогового) изображения в эквивалентный цифровой массив. В этой части книги рассматриваются операции дискретизации и квантования изображения, с помощью которых осуществляется такое преобразование. Описана также обратная операция, позволяющая из цифрового массива получить непрерывное изображение. Разработаны методы векторного представления изображений для детерминированных и случайных цифровых массивов. В заключение рассмотрен вопрос о качестве дискретизованных изображений, связанный с введением критериев верности их воспроизведения и дешифрируемости.

Глава 4

ДИСКРЕТИЗАЦИЯ И ВОССТАНОВЛЕНИЕ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ

В системы цифровой обработки изображений обычно поступают массивы чисел, полученные путем дискретизации реального изображения по пространственным переменным. После обработки образуются новые числовые массивы, используемые для восстановления непрерывного изображения, которое и рассматривает человек. Отсчеты изображения получаются в результате измерения некоторых физических характеристик реального изображения, как, например, яркости или оптической плотности. Любое измерительное устройство работает с какой-то погрешностью. Чтобы оценить достоверность измеренных значений и создать методы компенсации ошибок, важно иметь математическую модель ошибок измерения. Кроме того, часто не удается непосредственно измерить характеристики исходного изображения и вместо них измеряются некоторые величины, относящиеся к другому изображению, которое является функцией исходного. Для определения характеристик исходного изображения эту функцию приходится «обращать». Подобные операции обращения описаны в разделе, посвященном вопросам исправления (реставрации) изображений. В данной главе рассмотрены процессы дискретизации и восстановления непрерывных изображений применительно к идеальным и реальным системам.

4.1. ПРОЦЕССЫ ИДЕАЛЬНОЙ ДИСКРЕТИЗАЦИИ И ВОССТАНОВЛЕНИЯ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ

При разработке и анализе систем дискретизации и восстановления непрерывных изображений обрабатываемые изображения обычно принято рассматривать как детерминированные поля. Однако в некоторых случаях удобнее предполагать, что входной сигнал системы обработки изображений (особенно шумового происхождения) является реализацией двумерного случайного процесса. Ниже при анализе методов дискретизации и восстановления непрерывных изображений используются оба этих подхода.

4.1.1. ДИСКРЕТИЗАЦИЯ ДЕТЕРМИНИРОВАННЫХ ИЗОБРАЖЕНИЙ

Пусть функция $F_I(x, y)$ описывает исходное непрерывное изображение бесконечных размеров, представляя распределение яркости, оптической плотности или какого-либо другого параметра реального изображения. В идеальной системе дискретизации изображений пространственные отсчеты исходного изображения

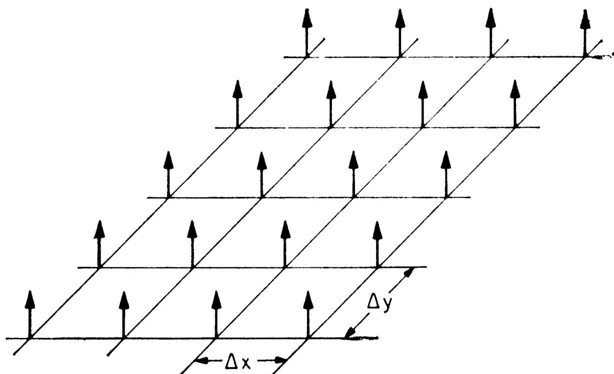


Рис. 4.1.1. Набор дельта-функций, осуществляющих дискретизацию изображений.

получаются фактически путем перемножения этой функции с пространственно-дискретизирующей функцией

$$S(x, y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(x - j_1 \Delta x,\ y - j_2 \Delta y), \tag{4.1.1}$$

состоящей из бесконечного числа дельта-функций, заданных в узлах решетки с шагом $(\Delta x, \Delta y)$, как показано на рис. 4.1.1. Тогда дискретизованное изображение описывается соотношением

$$F_P(x, y) = F_I(x, y)\, S(x, y) =$$
$$= \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} F_I(j_1 \Delta x,\ j_2 \Delta y)\, \delta(x - j_1 \Delta x,\ y - j_2 \Delta y), \tag{4.1.2}$$

в котором учитывается, что функцию $F_I(x, y)$ можно внести под знак суммирования и задать ее значения только в точках отсчета $(j_1 \Delta x,\ j_2 \Delta y)$. Для анализа процесса дискретизации удобно использовать спектр $\mathcal{F}_P(\omega_x, \omega_y)$, получаемый в результате непрерывного двумерного преобразования Фурье дискретизованного изображения

$$\mathcal{F}_P(\omega_x, \omega_y) = \iint_{-\infty}^{\infty} F_P(x, y) \exp\{-i(\omega_x x + \omega_y y)\}\, dx\, dy. \tag{4.1.3}$$

Согласно теореме о спектре свертки, спектр дискретизованного изображения можно представить в виде свертки спектра исходного изображения $\mathcal{F}_I(\omega_x, \omega_y)$ и спектра дискретизирующей функции $\mathcal{S}(\omega_x, \omega_y)$, т. е.

$$\mathcal{F}_P(\omega_x, \omega_y) = (1/4\pi^2)\mathcal{F}_I(\omega_x, \omega_y) * \mathcal{S}(\omega_x, \omega_y). \tag{4.1.4}$$

Двумерное преобразование Фурье дискретизирующей функции дает в результате бесконечный набор дельта-функций в плоскости пространственных частот с шагом $\omega_{xs} = 2\pi/\Delta x$ и $\omega_{ys} = 2\pi/\Delta y$ [4, стр. 22], т. е.

$$\mathcal{S}(\omega_x, \omega_y) = (4\pi^2/\Delta x\, \Delta y) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}). \tag{4.1.5}$$

Будем предполагать, что спектр исходного изображения ограничен по ширине так, что $\mathcal{F}_I(\omega_x, \omega_y) = 0$ при $|\omega_x| > \omega_{xc}$ и $|\omega_y| > \omega_{yc}$. Вычисляя свертку согласно равенству (4.1.4), найдем

$$\mathcal{F}_P(\omega_x, \omega_y) = (1/\Delta x\, \Delta y) \iint_{-\infty}^{\infty} \mathcal{F}_I(\omega_x - \alpha, \omega_y - \beta) \times \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(\alpha - j_1\omega_{xs}, \beta - j_2\omega_{ys})\, d\alpha\, d\beta. \tag{4.1.6}$$

Меняя порядок операций суммирования и интегрирования и учитывая основное свойство дельта-функции, получаем выражение для спектра дискретизованного изображения

$$\mathcal{F}_P(\omega_x, \omega_y) = (1/\Delta x\, \Delta y) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathcal{F}_I(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}). \tag{4.1.7}$$

Как показано на рис. 4.1.2, спектр дискретизованного изображения получается путем бесконечного повторения спектра исходного изображения со сдвигом на величины, кратные $(2\pi/\Delta x, 2\pi/\Delta y)$. Следует отметить, что если Δx и Δy выбраны слишком большими по сравнению с шириной спектра $\mathcal{F}_I(\omega_x, \omega_y)$, то соседние спектры будут перекрываться друг с другом.

Из отсчетов функции $F_P(x, y)$ можно получить непрерывное изображение путем линейной пространственной интерполяции или с помощью линейной пространственной фильтрации дискретизованного изображения. Пусть $R(x, y)$ есть импульсный отклик интерполирующего фильтра, а $\mathcal{R}(\omega_x, \omega_y)$ — его частотная характеристика. Восстановленное изображение получается как свертка последовательности отсчетов с импульсным откликом восстанавли-

вающего фильтра. Таким образсм, восстановленное непрерывное изображение описывается соотношением

$$F_R(x, y) = F_P(x, y) * R(x, y). \qquad (4.1.8)$$

Подставляя $F_P(x, y)$ из равенства (4.1.2) и вычисляя свертку, получаем

$$F_R(x, y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} F_I(j_1 \Delta x, j_2 \Delta y) R(x - j_1 \Delta x, y - j_2 \Delta y). \qquad (4.1.9)$$

Отсюда видно, что импульсный отклик $R(x, y)$ выполняет роль двумерной функции, интерполирующей отсчеты на всю плоскость. Пространственно-частотный спектр изображения, восстановленного согласно равенству (4.1.8), есть произведение частотной ха-

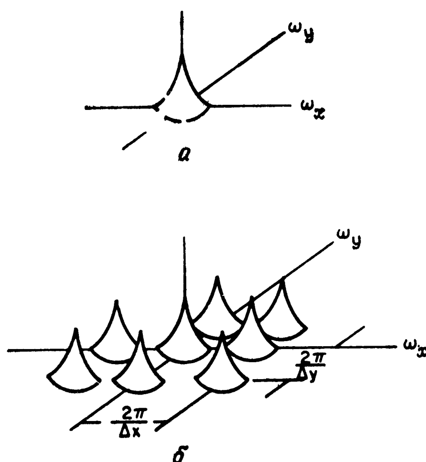


Рис. 4.1.2. Типичный спектр дискретизованного изображения.
а — исходное изображение; *б* — дискретизованное изображение.

рактеристики восстанавливающего фильтра со спектром дискретизованного изображения, т. е.

$$\mathcal{F}_R(\omega_x, \omega_y) = \mathcal{F}_P(\omega_x, \omega_y) \mathcal{R}(\omega_x, \omega_y). \qquad (4.1.10)$$

С учетом равенства (4.1.7) получаем

$$\mathcal{F}_R(\omega_x, \omega_y) = (1/\Delta x \Delta y) \mathcal{R}(\omega_x, \omega_y) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathcal{F}_I(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}). \qquad (4.1.11)$$

Из этого выражения видно, что если спектры не перекрываются, а множитель $\mathcal{R}(\omega_x, \omega_y)$ подавляет все сдвинутые спектры при j_1,

$j_2 \neq 0$, то спектр восстановленного непрерывного изображения будет совпадать со спектром исходного изображения и поэтому изображения также будут одинаковыми. Для изображений с ограниченной шириной спектра первое условие выполняется, если интервал дискретизации выбран так, что прямоугольная область, ограниченная верхними граничными частотами спектра изображения (ω_{xc}, ω_{yc}), лежит внутри прямоугольной области, определяемой половинами частот дискретизации ($\omega_{xs}/2$, $\omega_{ys}/2$). Следовательно, должны выполняться неравенства

$$\omega_{xc} \leqslant \omega_{xs}/2, \quad \omega_{yc} \leqslant \omega_{ys}/2, \tag{4.1.12а}$$

или

$$\Delta x \leqslant \pi/\omega_{xc}, \ \Delta y \leqslant \pi/\omega_{yc}. \tag{4.1.12б}$$

Это означает, что шаг дискретизации не должен превышать половины периода пространственной гармоники, соответствующей самым мелким деталям изображения. Это условие эквивалентно теореме о дискретизации одномерных сигналов, в которой сформулировано требование, что частота дискретизации должна хотя бы вдвое превышать наивысшую частоту спектра сигнала. Если соотношения (4.1.12) выполняются со знаками равенства, то изображение дискретизуется с найквистовской[1] частотой; если Δx и Δy меньше или больше, чем требуется по критерию Найквиста, то говорят, что изображение дискретизуется с избыточной или недостаточной частотой.

В тех случаях, когда пространственная частота дискретизации изображения достаточна для устранения наложения спектров в дискретизованном изображении, исходное изображение можно абсолютно точно восстановить путем пространственной фильтрации отсчетов с помощью соответствующего фильтра. Так, например, фильтр, частотная характеристика которого приведена на рис. 4.1.3 и описывается выражениями

$$\mathcal{R}(\omega_x, \omega_y) = K \text{ при } |\omega_x| \leqslant \omega_{xL} \text{ и } |\omega_y| \leqslant \omega_{yL}, \tag{4.1.13а}$$

$$\mathcal{R}(\omega_x, \omega_y) = 0 \text{ в остальных случаях}, \tag{4.1.13б}$$

где K — масштабная постоянная, удовлетворяет условию точного восстановления, если $\omega_{xL} > \omega_{xc}$ и $\omega_{yL} > \omega_{yc}$. Функция рассеяния точки (или импульсный отклик) данного восстанавливающего фильтра имеет вид

$$R(x, y) = (K\omega_{xL}\omega_{yL}/\pi^2) [\sin(\omega_{xL}x)/(\omega_{xL}x)] [\sin(\omega_{yL}y)/(\omega_{yL}y)]. \tag{4.1.14}$$

При использовании этого фильтра изображение восстанавливается с помощью бесконечной суммы функций вида sin c (x). Другим

[1] В отечественной литературе это положение известно как теорема Котельникова. — *Прим. перев.*

фильтром, пригодным для восстановления изображений, является «круговой» фильтр с частотной характеристикой (рис. 4, 1, 3, *б*)

$$\mathcal{R}(\omega_x,\ \omega_y) = K \text{ при } \sqrt{\omega_x^2 + \omega_y^2} \leqslant \omega_0, \qquad (4.1.15а)$$

$$\mathcal{R}(\omega_x,\ \omega_y) = 0 \text{ в остальных случаях}, \qquad (4.1.15б)$$

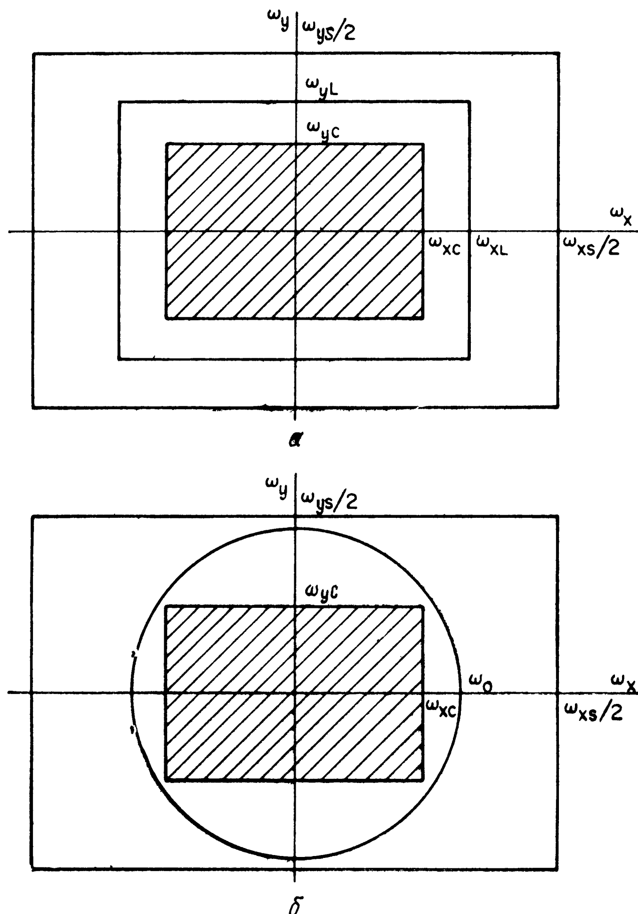


Рис. 4.1.3. Фильтры для восстановления дискретизованных изображений.
а — прямоугольный; *б* — круговой.

если $\omega_0^2 < \omega_{xc}^2 + \omega_{yc}^2$. Импульсный отклик такого фильтра имеет вид

$$R(x,\ y) = 2\pi\omega_0 \left(J_1 \{\omega_0 \sqrt{x^2 + y^2}\} / \sqrt{x^2 + y^2}\right), \qquad (4.1.16)$$

где $J_1(x)$ — бесселева функция первого порядка. Существует много восстанавливающих фильтров (или соответствующих интер-

поляционных функций), которые можно использовать для восстановления изображений. Однако на практике в цифровой системе воспроизведения изображений часто бывает сложно реализовать оптимальный восстанавливающий фильтр. Одна из практических трудностей состоит в том, что обычные интерполяционные функции [например, заданные равенствами (4.1.14) и (4.1.16)] принимают не только положительные, но и отрицательные значения, хотя функции, описывающие восстановленные изображения, строго положительны. Такие интерполяционные функции нельзя сформировать оптическими средствами и, следовательно, не удается восстановить изображения путем последовательного их взвешивания и сложения. Ричардс [6] нашел семейство интерполяционных функций, которые можно применить для последовательной оптической интерполяции, поскольку их большие значения положительны.

4.1.2. ДИСКРЕТИЗАЦИЯ СЛУЧАЙНЫХ ИЗОБРАЖЕНИЙ [5, 7, 8]

Выше при анализе вопросов о дискретизации и восстановлении изображений предполагалось, что функция $F_I(x, y)$, описывающая исходное изображение, является детерминированной. Было показано, что если это изображение имеет спектр ограниченной ширины, то совокупность отсчетов изображения, взятых с найквистовской частотой в дискретных точках, достаточна, чтобы с помощью интерполяции восстановить из отсчетов точную копию исходного изображения. Подобный вывод справедлив и в случае дискретизации случайных двумерных полей. Пусть $F_I(x, y)$ обозначает непрерывный двумерный стационарный случайный процесс с известными средним значением η_{F_I} и автокорреляционной функцией

$$R_{F_I}(\tau_x, \tau_y) = E\{F_I(x_1, y_1) F_I^*(x_2, y_2)\}, \tag{4.1.17}$$

где $\tau_x = x_1 - x_2$ и $\tau_y = y_1 - y_2$. С помощью набора дельта-функций произведем дискретизацию этого процесса. В результате получим

$$F_P(x, y) = F_I(x, y) S(x, y) =$$
$$= F_I(x, y) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(x - j_1 \Delta x, y - j_2 \Delta y). \tag{4.1.18}$$

Автокорреляционная функция дискретизованного процесса имеет вид

$$R_{F_P}(x_1, x_2; \; y_1, y_2) = E\{F_P(x_1, y_1) F_P^*(x_2, y_2)\} =$$
$$= E\{F_I(x_1, y_1) F_I^*(x_2, y_2)\} S(x_1, y_1) S(x_2, y_2). \tag{4.1.19}$$

Первый сомножитель правой части этого равенства является автокорреляционной функцией стационарного исходного изобра-

жения. Следует учесть, что стоящее в правой части равенства (4.1.19) произведение двух дискретизирующих дельта-функций есть дельта-функция вида

$$S(x_1, y_1) S(x_2, y_2) = S(x_1 - x_2, y_1 - y_2) = = S(\tau_x, \tau_y). \qquad (4.1.20)$$

Следовательно, дискретизованное случайное поле, представляющее изображение, также стационарно и имеет автокорреляционную функцию

$$R_{F_P}(\tau_x, \tau_y) = R_{F_I}(\tau_x, \tau_y) S(\tau_x, \tau_y). \qquad (4.1.21)$$

С помощью двумерного преобразования Фурье автокорреляционной функции (4.1.21) найдем энергетический спектр дискретизованного случайного поля. Согласно теореме о спектре произведения,

$$\mathcal{W}_{F_P}(\omega_x, \omega_y) = (1/4\pi^2)\mathcal{W}_{F_I}(\omega_x, \omega_y) * \mathcal{S}(\omega_x, \omega_y), \qquad (4.1.22)$$

где $\mathcal{W}_{F_I}(\omega_x, \omega_y)$ и $\mathcal{W}_{F_P}(\omega_x, \omega_y)$ обозначают соответственно спектральные плотности исходного и дискретизованного изображений, а $\mathcal{S}(\omega_x, \omega_y)$ — результат преобразования Фурье дискретизирующих дельта-функций. Далее, повторяя шаги, сделанные при выводе соотношения (4.1.7), получим, что энергетический спектр дискретизованного изображения можно записать в следующем виде:

$$\mathcal{W}_{F_P}(\omega_x, \omega_y) = (4\pi^2/\Delta x\, \Delta y) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathcal{W}_{F_I}(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}). \qquad (4.1.23)$$

Таким образом, энергетический спектр дискретизованного изображения образуется повторением энергетического спектра непрерывного исходного изображения в пространственно-частотной области через интервалы, кратные пространственной частоте дискретизации $(2\pi/\Delta x, 2\pi/\Delta y)$. Если ширина энергетического спектра непрерывного изображения ограничена так, что $\mathcal{W}_{F_I}(\omega_x, \omega_y) = 0$ при $|\omega_x| > \omega_{xc}$ и $|\omega_y| > \omega_{yc}$, где ω_{xc} и ω_{yc} — граничные частоты, то в сумме (4.1.23) отдельные спектры не будут накладываться в том случае, когда интервалы дискретизации удовлетворяют условиям $\Delta x < \pi/\omega_{xc}$ и $\Delta y < \pi/\omega_{yc}$. Непрерывное случайное изображение $F_R(x, y)$ можно восстановить из отсчетов исходного случайного изображения с помощью интерполяции по формуле

$$F_R(x, y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} F_I(j_1\Delta x, j_2\Delta y) R(x - j_1\Delta x, y - j_2\Delta y), \qquad (4.1.24)$$

где $R(x, y)$ — детерминированная интерполяционная функция. Можно добиться, чтобы восстановленное поле и исходное изображение были эквивалентны в среднеквадратическом смысле [5, стр. 284], т. е.

$$E\{|F_I(x, y) - F_R(x, y)|^2\} = 0. \tag{4.1.25}$$

Для этого достаточно, чтобы удовлетворялся критерий Найквиста и в качестве интерполяционной функции была выбрана достаточно

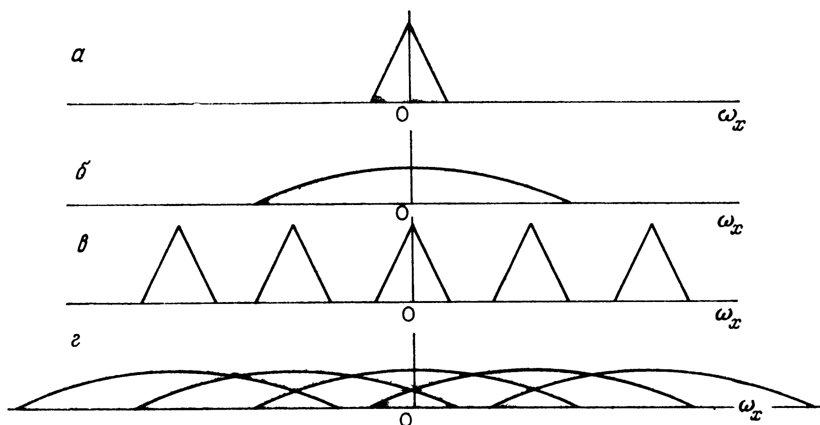


Рис. 4.1.4. Спектры дискретизованного изображения, содержащего сигнал и помеху.
а — сигнал; *б* — помеха; *в* — дискретизованный сигнал; *г* — дискретизованная помеха.

«хорошая» функция, как, например, бесселева или функция sinc (x), определяемые формулами (4.1.16) и (4.1.14).

Приведенные результаты можно непосредственно применить к практической задаче о дискретизации суммы детерминированного изображения и аддитивного шума, моделью которого служит случайное поле. На рис. 4.1.4 показан пример энергетического спектра дискретизованного зашумленного изображения. Этот пример указывает на возможность возникновения серьезной трудности. Спектр шума может оказаться шире спектра полезного изображения, и если шум будет дискретизироваться с недостаточной частотой, то в полосу пропускания восстанавливающего фильтра будут попадать «хвосты» шумового спектра. Это приведет к увеличению искажений, вызванных шумом. Для устранения этой трудности перед дискретизацией следует, конечно, профильтровать зашумленное изображение с тем, чтобы сузить спектр шума.

4.2. РЕАЛЬНЫЕ СИСТЕМЫ ДИСКРЕТИЗАЦИИ ИЗОБРАЖЕНИЙ

В реальной системе дискретизации изображений дискретизирующая решетка имеет конечные размеры, ширина дискретизирующих импульсов заметно отличается от нуля, а отсчеты изображения могут быть взяты с недостаточной частотой. Рассмотрим последствия такой неидеальной дискретизации.

Основой для анализа будет служить типичная система сканирования изображений, показанная на рис. 4.2.1. Исходное изображение, представленное в виде диапозитива, сканируется скачкообразно перемещающимся тонким световым лучом. Свет, прошедший через диапозитив, собирается конденсорной линзой и направляется на светочувствительную поверхность фотодетектора. Электрический сигнал с выхода фотодетектора интегрируется за период времени, в течение которого световой луч освещает данный элемент изображения. В ходе анализа будем предполагать, что в процессе дискретизации дополнительных шумов в сигнал не

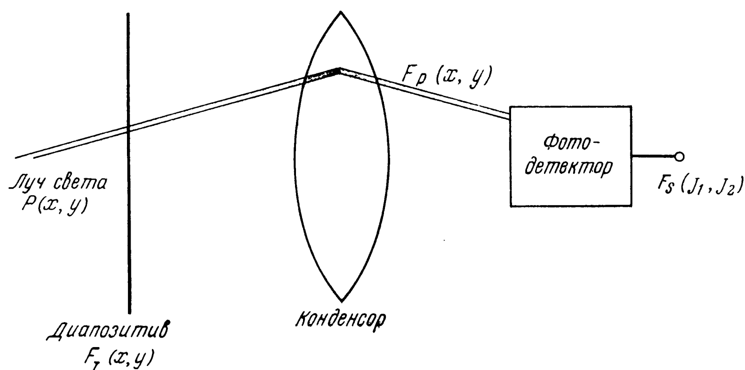


Рис. 4.2.1. Типичная система сканирования изображений.

вносится. Поэтому результаты предыдущего раздела, относящиеся к дискретизации зашумленных изображений, нетрудно объединить с выводами данного раздела. Следует также отметить, что нижеприведенный анализ легко распространяется на широкий класс реальных систем дискретизации изображений.

4.2.1. ВЛИЯНИЕ ФОРМЫ ДИСКРЕТИЗИРУЮЩЕГО ИМПУЛЬСА

С учетом вышеупомянутых предположений дискретизованное изображение можно описать функцией

$$F_P(x, y) = F_I(x, y)\, S(x, y), \tag{4.2.1}$$

где дискретизирующая функция

$$S(x, y) = \sum_{j_1=-J_1}^{J_1} \sum_{j_2=-J_2}^{J_2} P(x - j_1 \Delta x, y - j_2 \Delta y) \qquad (4.2.2)$$

состоит из $(2J_1 + 1)(2J_2 + 1)$ одинаковых импульсов $P(x, y)$, образующих решетку с шагом $(\Delta x, \Delta y)$. Для упрощения обозначений пределы суммирования выбраны симметричными. Будем считать, что дискретизирующие импульсы нормированы так, что

$$\iint_{-\infty}^{\infty} P(x, y)\, dx\, dy = 1. \qquad (4.2.3)$$

При анализе можно полагать, что дискретизирующая функция была получена пропусканием конечного набора дельта-функций $D_T(x, y)$ через линейный фильтр с импульсным откликом $P(x, y)$. Таким образом,

$$S(x, y) = D_T(x, y) * P(x, y), \qquad (4.2.4)$$

где

$$D_T(x, y) = \sum_{j_1=-J_1}^{J_1} \sum_{j_2=-J_2}^{J_2} \delta(x - j_1 \Delta x, y - j_2 \Delta y). \qquad (4.2.5)$$

Подставив выражение (4.2.2) в (4.2.1), получим формулу для дискретизованного изображения

$$F_P(x, y) = \sum_{j_1=-J_1}^{J_1} \sum_{j_2=-J_2}^{J_2} F_I(x, y) P(x - j_1 \Delta x, y - j_2 \Delta y). \qquad (4.2.6)$$

Спектр этой функции имеет вид

$$\mathcal{F}_P(\omega_x, \omega_y) = (1/4\pi^2)\mathcal{F}_I(\omega_x, \omega_y) * [\mathcal{D}_T(\omega_x, \omega_y)\mathcal{P}(\omega_x, \omega_y)], \qquad (4.2.7)$$

где $\mathcal{P}(\omega_x, \omega_y)$ — результат преобразования Фурье функции $P(x, y)$. Преобразование Фурье конечной решетки дискретизирующих импульсов описывается следующим соотношением [5, стр. 105]:

$$\mathcal{D}_T(\omega_x, \omega_y) = \frac{\sin\left[\omega_x\left(J_1 + \frac{1}{2}\right)\Delta x\right]}{\sin(\omega_x \Delta x/2)} \frac{\sin\left[\omega_y\left(J_2 + \frac{1}{2}\right)\Delta y\right]}{\sin(\omega_y \Delta y/2)}. \qquad (4.2.8)$$

На рис. 4.2.2 приведен график функции $\mathcal{D}_T(\omega_x, \omega_y)$. С увеличением J_1 и J_2 правая часть формулы (4.2.8) в пределе превращается в набор дельта-функций.

В системе восстановления изображений непрерывное изображение получается путем интерполяции отсчетов. Идеальные интерполяционные функции, такие, как $\operatorname{sinc}(x)$ и бесселевы [формулы (4.1.14) и (4.1.16)], обычно определены на бесконечной плоскости. Если же дискретизирующая решетка имеет конечные

размеры, то на границах отсекаются «хвосты» интерполяционных функций и вблизи краев восстановленного изображения появляются ошибки [9, 10]. Однако такие ошибки обычно становятся пренебрежимо малыми при удалении от границ на 8—10 шагов дискретизации.

Числовые значения отсчетов изображения получаются путем пространственного интегрирования $F_S(x, y)$ по некоторой конечной площадке — элементу изображения. В сканирующей системе

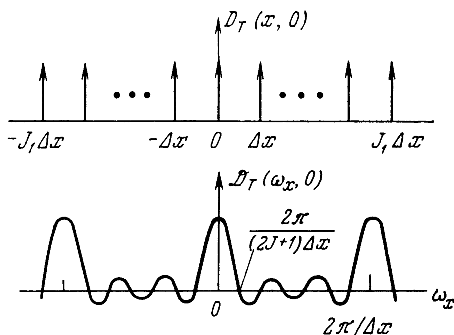


Рис. 4.2.2. Усеченная дискретизирующая последовательность и ее спектр.

(рис. 4.2.1) интегрирование фактически проводится на светочувствительной поверхности фотодетектора. Значение отсчета, соответствующего (j_1, j_2)-му элементу, можно найти по формуле

$$F_S(j_1\Delta x, j_2\Delta y) = \int_{j_1\Delta x - A_x}^{j_1\Delta x + A_x} \int_{j_2\Delta y - A_y}^{j_2\Delta y + A_y} F_I(x, y) P(x - j_1\Delta x, y - j_2\Delta y)\, dx\, dy, \quad (4.2.9)$$

где A_x и A_y обозначают наибольшие размеры этого элемента. Здесь предполагается, что за время интегрирования в системе берется лишь один отсчет. В противном случае приходится решать сложную проблему перекрестных искажений. В рассматриваемой системе дискретизации размеры элемента могут оказаться больше, чем расстояние между отсчетами. Поэтому в модели допускается, что последовательные (во времени) отсчеты соответствуют частично перекрывающимся элементам изображения.

Простой заменой переменных равенство (4.2.9) можно преобразовать к виду

$$F_S(j_1\Delta x, j_2\Delta y) = \int_{-A_x}^{A_x} \int_{-A_y}^{A_y} F_I(j_1\Delta x - \alpha, j_2\Delta y - \beta) P(-\alpha, -\beta)\, d\alpha\, d\beta. \quad (4.2.10)$$

Поскольку предполагается, что за время интегрирования берется только один отсчет, пределы в интеграле (4.2.10) можно расширить до бесконечности. В такой форме выражение (4.2.10) можно рассматривать как результат свертки исходного изображения $F_I(x, y)$ с импульсным откликом $P(-x, -y)$ и последующей дискретизации этой свертки в конечной области с помощью дельта-функции. Тогда, пренебрегая эффектами, связанными с конечными размерами дискретизирующей решетки дельта-функций, получим

$$F_S(j_1 \Delta x, j_2 \Delta y) = [F_I(x, y) \circledast P(-x, -y)]\, \delta(x - j_1 \Delta x, y - j_2 \Delta y). \tag{4.2.11}$$

В большинстве систем дискретизации дискретизирующий импульс симметричен, поэтому $P(-x, -y) = P(x, y)$.

Несложное по форме соотношение (4.2.11) полезно при оценке эффектов, возникающих при дискретизации с использованием импульсов конечной ширины. Если спектр изображения ограничен по ширине, а Δx и Δy удовлетворяют критерию Найквиста, то конечная ширина импульса приводит к тем же результатам, как если бы исходное изображение перед идеальной дискретизацией подверглось линейному искажению (смазыванию). В части 4 будут рассмотрены методы компенсации подобных искажений. Однако конечность размеров дискретизирующего импульса не всегда является недостатком. Рассмотрим случай, когда спектр исходного изображения очень широкий, и поэтому оно дискретизируется с недостаточной частотой. Импульс с конечными размерами фактически осуществляет низкочастотную фильтрацию исходного изображения, что приводит к сужению спектра и, следовательно, уменьшает ошибки, вызванные наложением спектров.

4.2.2. ЭФФЕКТЫ, СВЯЗАННЫЕ С НАЛОЖЕНИЕМ СПЕКТРОВ

Для безошибочного восстановления непрерывного изображения необходимо ограничить ширину спектра исходного изображения, взять отсчеты по пространственным переменным с частотой не ниже найквистовской и интерполировать отсчеты изображения. Вопрос об интерполяции рассмотрен в следующем разделе, а здесь дан анализ эффектов, возникающих при дискретизации с недостаточной частотой.

Если при такой дискретизации имеет место наложение спектров, как отмечено штриховкой на рис. 4.2.3, то в восстановленное изображение будут внесены ложные пространственные гармоники [11, 12]. На рис. 4.2.4 показано, как выглядят на реальном изображении ошибки, вызванные наложением спектров. Недостаточная частота дискретизации приводит к появлению в восстановленном изображении ложных низкочастотных гармоник. В оптике это явление называют муаровыми узорами или муар-эффектом.

Согласно равенству (4.1.7), спектр дискретизованного изображения можно представить в виде

$$\mathcal{F}_P(\omega_x, \omega_y) = (1/\Delta x\, \Delta y)[\mathcal{F}_I(\omega_x, \omega_y) + \mathcal{F}_Q(\omega_x, \omega_y)], \quad (4.2.12)$$

где $\mathcal{F}_I(\omega_x, \omega_y)$ описывает спектр исходного изображения, дискретизованного с периодом $(\Delta x, \Delta y)$. Слагаемое

$$\mathcal{F}_Q(\omega_x, \omega_y) = \underbrace{\sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty}}_{\substack{j_1 \neq 0 \\ j_2 \neq 0}} \mathcal{F}_I(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}) \quad (4.2.13)$$

описывает те компоненты спектра дискретизованного изображения, которые получаются при повторении исходного спектра с перио-

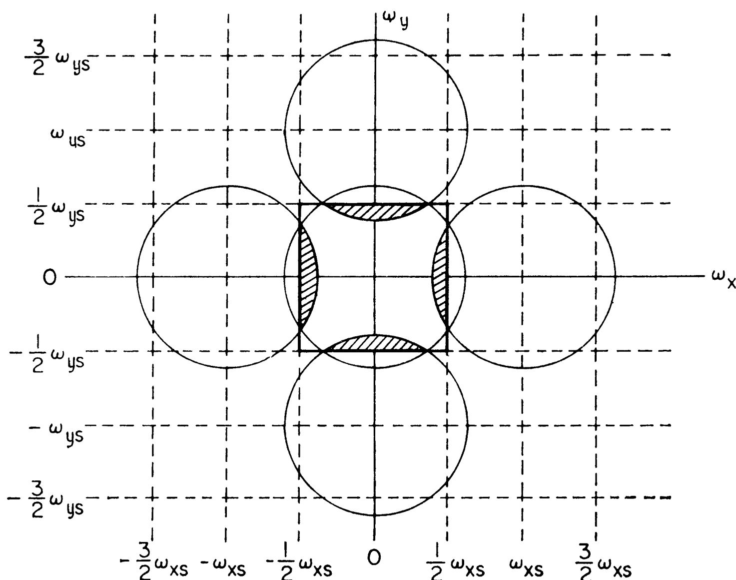
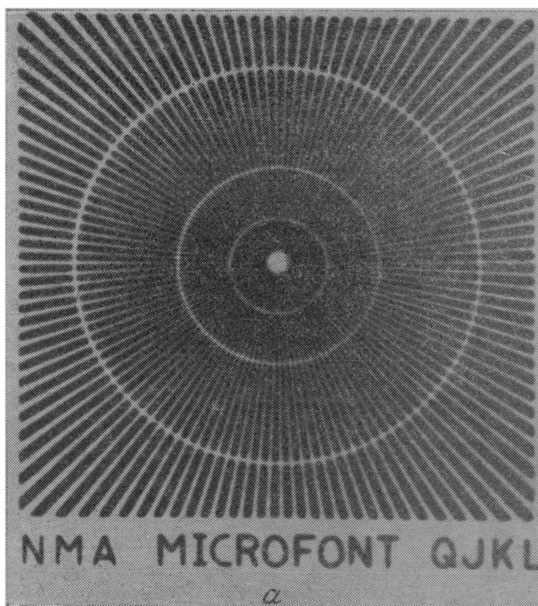


Рис. 4.2.3. Спектры двумерной функции, дискретизованной с недостаточной частотой.

дами $\omega_{xs} = 2\pi/\Delta x$ и $\omega_{ys} = 2\pi/\Delta y$. При отсутствии наложения спектров оптимальную интерполяцию отсчетов изображения можно получить, пропуская дискретизованное изображение через идеальный фильтр нижних частот с частотной характеристикой

$$\mathcal{R}(\omega_x, \omega_y) = \begin{cases} K \text{ при } |\omega_x| \leqslant \omega_{xs}/2 \text{ и } |\omega_y| \leqslant \omega_{ys}/2, & (4.2.14а) \\ 0 \text{ в остальных случаях}, & (4.2.14б) \end{cases}$$

где K — нормирующая постоянная. Если этот же метод применить для восстановления изображения, дискретизованного с недоста-



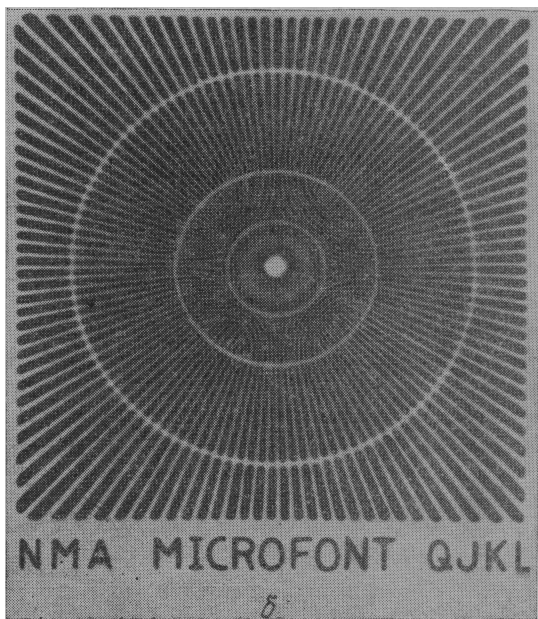


Рис. 4.2.4. Пример искажений, вызванных наложением спектров.
а — исходное изображение; *б* — дискретизованное изображение.

точной частотой, то получится изображение, описываемое функцией

$$F_R(x, y) = F_I(x, y) + A(x, y), \tag{4.2.15}$$

где слагаемое

$$A(x, y) = (1/4\pi^2) \int_{-\omega_{xs}/2}^{\omega_{xs}/2} \int_{-\omega_{ys}/2}^{\omega_{ys}/2} \mathscr{F}_Q(\omega_x, \omega_y) \times \exp\{i[\omega_x x + \omega_y y]\} \, d\omega_x \, d\omega_y \tag{4.2.16}$$

соответствует тем деталям восстановленного изображения, которые появились вследствие наложения. Все коэффициенты, изме-

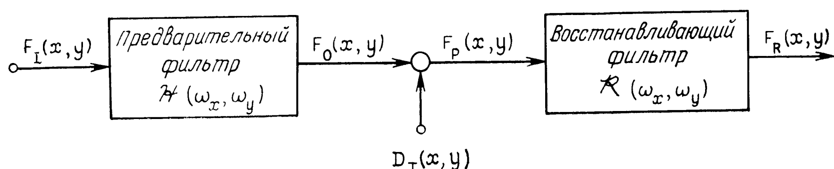


Рис. 4.2.5. К анализу ошибок, вызванных наложением спектров.

няющие амплитуду, входят в K. Ложную компоненту (4.2.16) можно существенно ослабить, если для интерполяции применить фильтр, частотная характеристика которого согласована со спектром исходного изображения. Этот метод описан в гл. 15. Ниже будет рассмотрен другой подход, при котором для уменьшения наложения спектров изображение перед дискретизацией подвергают низкочастотной фильтрации.

На рис. 4.2.5 представлена модель, служащая для количественного анализа эффектов, связанных с наложением спектров. В ней предполагается, что исходное изображение $F_I(x, y)$ является реализацией двумерного случайного процесса с известной спектральной плотностью $\mathscr{F}_I(\omega_x, \omega_y)$. Исходное изображение пропускают через предварительный линейный пространственный фильтр с частотной характеристикой $\mathscr{H}(\omega_x, \omega_y)$. Будем полагать, что этот фильтр обеспечивает плавное ослабление высших пространственных гармоник (т. е. это не идеальный фильтр нижних частот с резкой отсечкой). Затем отфильтрованное изображение дискретизируется с шагом $(\Delta x, \Delta y)$ с помощью идеального пространственного дискретизатора. Далее отсчеты интерполируются в восстанавливающем фильтре и получается копия исходного изображения. Согласно равенству (1.10.8), спектральная плотность изображения на выходе предварительного фильтра равна

$$\mathscr{W}_{\mathscr{F}_0}(\omega_x, \omega_y) = |\mathscr{H}(\omega_x, \omega_y)|^2 \mathscr{W}_{\mathscr{F}_I}(\omega_x, \omega_y), \tag{4.2.17}$$

а спектр дискретизованного изображения

$$\mathscr{W}_{\mathscr{F}_P}(\omega_x, \omega_y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathscr{W}_{\mathscr{F}_0}(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}). \tag{4.2.18}$$

Желательно разделить эффекты, вызванные дискретизацией с недостаточной частотой, и последствия неточного восстановления. Поэтому здесь будем полагать, что в качестве восстанавливающего фильтра взят оптимальный фильтр с частотной характеристикой (4.2.14). Мощность компоненты сигнала при $j_1 = j_2 = 0$, проходящей через такой фильтр, при этом равна

$$E_R = \int_{-\omega_{xs}/2}^{\omega_{xs}/2} \int_{-\omega_{ys}/2}^{\omega_{ys}/2} \mathscr{W}_{\mathscr{F}_I}(\omega_x, \omega_y) \,|\mathscr{H}(\omega_x, \omega_y)|^2 \,d\omega_x \,d\omega_y. \tag{4.2.19}$$

В идеальном случае частотная характеристика предварительного фильтра должна совпадать с частотной характеристикой (4.2.14) восстанавливающего фильтра. При этом мощность дискретизованного сигнала оказывается наибольшей и равна

$$E_{RM} = \int_{-\omega_{xs}/2}^{\omega_{xs}/2} \int_{-\omega_{ys}/2}^{\omega_{ys}/2} \mathscr{W}_{\mathscr{F}_I}(\omega_x, \omega_y) \,d\omega_x \,d\omega_y. \tag{4.2.20}$$

Уменьшение резкости изображения, связанное с введением предварительного фильтра, можно выразить отношением

$$\mathscr{E}_R = (E_{RM} - E_R)/E_{RM}. \tag{4.2.21}$$

Ошибку, вызванную наложением спектров, обычно определяют через энергию паразитных гармоник, принадлежащих смещенным спектрам, попадающим в полосу пропускания восстанавливающего фильтра. Предположим для упрощения, что частота дискретизации довольно высока и вклад спектров, смещенных на $(\pm j_1\omega_{xs}/2, \ \pm j_2\omega_{ys}/2)$, пренебрежимо мал при $j_1 \geqslant 2$ и $j_2 \geqslant 2$. Полная энергия ошибки, вызванной наложением спектров (см. заштрихованные участки на рис. 4.2.3), равна

$$E_A = E_0 - E_R, \tag{4.2.22}$$

где

$$E_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathscr{W}_{\mathscr{F}_I}(\omega_x, \omega_y) \,|\mathscr{H}(\omega_x, \omega_y)|^2 \,d\omega_x \,d\omega_y \tag{4.2.23}$$

есть энергия сигнала на выходе предварительного фильтра. Тогда ошибку из-за наложения спектров можно определить как [11]

$$\mathscr{E}_A = E_A/E_0. \tag{4.2.24}$$

Эту ошибку можно уменьшить путем ослабления высших пространственных гармоник $F_I(x, y)$ в предварительном фильтре. Однако любое ослабление гармоник, попадающих в полосу пропускания восстанавливающего фильтра, приводит к ухудшению резкости изображения. Таким образом, приходится идти на компромисс между резкостью дискретизованного изображения и ошибкой, вызванной наложением спектров.

Рассмотрим теперь соотношение между величиной этой ошибки и степенью снижения резкости применительно к нескольким типам предварительных фильтров. Вероятно, простейшим способом пространственной фильтрации изображения, полученного в некогерентном свете, является пропускание изображения через линзу с ограниченной апертурой. Пространственную фильтрацию также можно осуществить, дефокусируя оптическую систему. На рис. 13.2.2 показана зависимость частотной характеристики круглой линзы от степени ее дефокусировки. Даже при использовании точно сфокусированной линзы, не имеющей аберраций, получается некоторое смазывание изображения из-за дифракции на конечной апертуре. Частотная характеристика безаберрационной круглой линзы диаметром d, вносящей только дифракционные искажения, равна [13, стр. 83]

$$\mathcal{H}(\omega) = (2/\pi)\left[\arccos(\omega/\omega_0) - (\omega/\omega_0)\sqrt{1-(\omega/\omega_0)^2}\right], \quad 0 \leqslant |\omega| \leqslant \omega_0, \tag{4.2.25а}$$

$$\mathcal{H}(\omega) = 0, \quad |\omega| > \omega_0, \tag{4.2.25б}$$

где $\omega_0 = \pi d/R$, а R — расстояние от линзы до фокальной плоскости. В предыдущем разделе было отмечено, что неидеальная дискретизация, когда ширина дискретизирующего импульса конечна, эквивалентна идеальной дискретизации того же изображения, но пропущенного через пространственный фильтр, импульсный отклик которого (после изменения направления обеих координатных осей) совпадает с формой дискретизирующего импульса. Таким образом, дискретизирующий импульс может производить предварительную фильтрацию.

В некогерентной оптической системе сканирующего микроденситометра дискретизирующий импульс имеет прямоугольную форму:

$$P(x, y) = \begin{cases} 1/T^2, & |x, y| \leqslant T/2, \\ 0, & |x, y| > T/2. \end{cases} \tag{4.2.26}$$

Частотная характеристика в этом случае равна

$$\mathcal{P}(\omega_x, \omega_y) = \frac{\sin(\omega_x T/2)}{(\omega_x T/2)} \frac{\sin(\omega_y T/2)}{(\omega_y T/2)}. \tag{4.2.27}$$

В электронно-лучевых трубках образуется световое пятно с двумерным гауссовым распределением яркости

$$P(x,\ y)=\frac{1}{2\pi\sigma_w^2}\exp\left\{-\frac{(x^2+y^2)}{2\sigma_w^2}\right\},\tag{4.2.28}$$

где σ_w характеризует ширину луча. Частотная характеристика сканирующей системы с гауссовым пятном имеет вид

$$\mathscr{P}(\omega_x,\ \omega_y)=\exp\left\{-\frac{(\omega_x^2+\omega_y^2)\,\sigma_w^2}{2}\right\}.\tag{4.2.29}$$

На рис. 4.2.6 приведены графики зависимости потери резкости и величины ошибки, вызванной наложением спектров, от ω_0 для безаберрационной линзы, а также для квадратного и гауссова пятен. Ширина квадратного пятна выбрана равной $T=2\pi/\omega_0$,

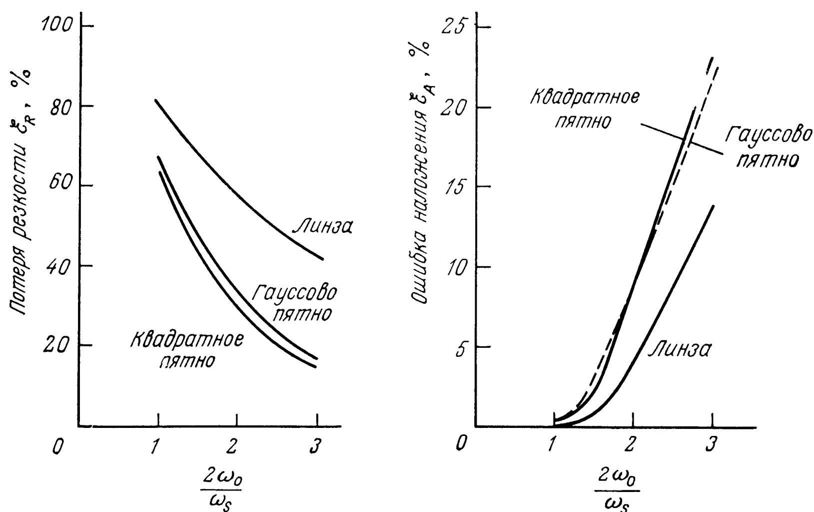


Рис. 4.2.6. Ошибка, вызванная наложением спектров, и потеря резкости для различных предварительных фильтров при $m=1$.

так что первый нуль функции sinc (x) совпадает с частотой отсечки частотной характеристики линзы. Рассеяние гауссова пятна выбрано равным $\sigma_w=2/\omega_0$. В данном примере спектр исходного изображения моделировался функцией

$$\mathscr{W}_{\mathscr{F}_I}(\omega_x,\ \omega_y)=A/[1+(\omega/\omega_c)^{2m}],\tag{4.2.30}$$

где A — амплитудный множитель, m — целочисленный параметр, определяющий скорость, с которой амплитуда спектра уменьшается с ростом частоты, а ω_c — пространственная частота, на

которой амплитуда спектра уменьшается в два раза. Кривые, приведенные на рис. 4.2.6, показывают, что предварительная фильтрация, осуществляемая при гауссовой и квадратной формах сканирующего пятна, дает примерно одинаковые результаты, а безаберрационная линза с конечной апертурой при той же величине ошибки, вызванной наложением спектров, дает несколько большую потерю резкости. При использовании дефокусированной линзы получаются самые плохие результаты.

4.3. РЕАЛЬНЫЕ СИСТЕМЫ ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЙ

В разд. 4.1 были сформулированы условия, необходимые для точного восстановления изображения. Из исходного изображения нужно взять отсчеты с частотой, вдвое превышающей наивысшую пространственную частоту его спектра. Кроме того, восстанавливающий фильтр (или эквивалентный интерполятор) должен быть рассчитан так, чтобы пропускал без искажений основной спектр при $j_1 = 0$, $j_2 = 0$ и полностью подавлял все побочные спектры при j_1, $j_2 \neq 0$. Ниже приведены некоторые соображения о том, к каким результатам может привести применение неидеальных восстанавливающих фильтров.

В большинстве систем цифровой обработки изображений электрические сигналы, представляющие собой отсчеты изображения, поступают с выхода вычислительной машины последовательно, согласно обычной схеме развертки. Эти «электрические отсчеты» следует подать на устройство воспроизведения (синтеза) непрерывных изображений, например на электронно-лучевую трубку (ЭЛТ) или фоторегистратор. Ниже рассмотрены три способа выполнения такой операции.

4.3.1. СПОСОБЫ ВОСПРОИЗВЕДЕНИЯ ИЗОБРАЖЕНИЙ

На рис. 4.3.1 изображена схема некогерентной оптической системы восстановления изображений. Маленькое световое пятно скачками перемещается по экрану ЭЛТ, образуя растр; яркость каждой точки модулируется пропорционально значениям отсчетов изображения. С помощью проекционной системы, содержащей некогерентный пространственный фильтр с желательными свойствами, картинку с экрана ЭЛТ можно затем отобразить на большой экран для просмотра или на фотопленку для записи. Как правило, оптимальный фильтр, равномерно пропускающий спектр изображения и имеющий резкую отсечку за пределами спектра, физически реализовать не удается.

Для восстановления одноцветных изображений можно применить когерентную оптическую систему, схема которой приведена на рис. 4.3.2. В этой системе сначала фотографируют картинку

с экрана ЭЛТ и изготавливают диапозитив, который освещают коллимированным лазерным лучом. Свет, прошедший через диапозитив, попадает на линзу, которая в своей задней фокальной плоскости создает световое поле с распределением интенсивности, пропорциональным двумерному спектру Фурье пространствен-

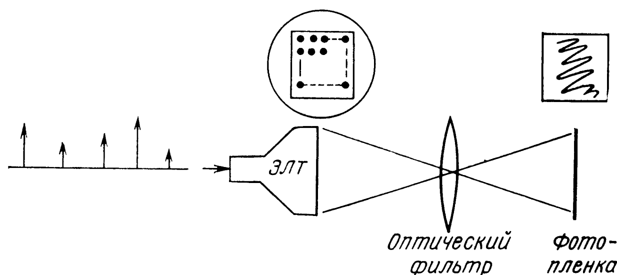


Рис. 4.3.1. Некогерентная оптическая система восстановления изображений.

ного распределения коэффициента пропускания диапозитива. Можно рассчитать оптический фильтр, который желаемым образом изменял бы распределение амплитуды и фазы светового поля в плоскости фильтрации, т. е. мог бы играть роль восстанавливающего фильтра. Вторая линза также выполняет преобразование Фурье и восстанавливает изображение на поверхности фотопленки.

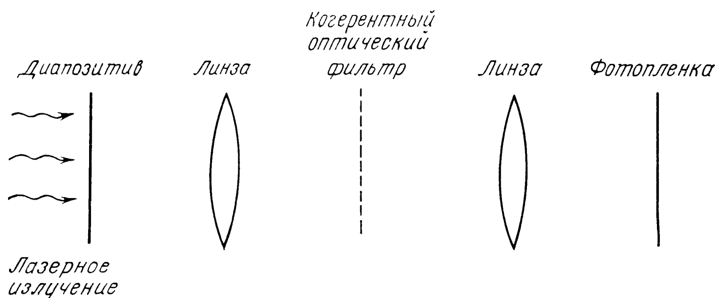


Рис. 4.3.2. Когерентная оптическая система восстановления изображений.

Главным достоинством устройства, построенного на основе когерентной оптической системы, является сравнительная простота изготовления восстанавливающего фильтра. Оптимальный фильтр является просто диафрагмой, пропускающей только дифракционную картинку нулевого порядка.

В системах восстановления изображений широко используются электронно-оптические устройства. Так, например, можно весьма просто осуществить интерполяцию путем электрической дефоку-

сировки луча на экране ЭЛТ. Недостатки такого метода состоят в том, что трудно поддерживать на всем экране заданную форму светового пятна, а интерполяция, как правило, оказывается субоптимальной. В сканирующих фоторегистраторах обычно восстанавливают непрерывное изображение, проектируя прямоугольное световое пятно на фотопленку. В большинстве случаев размер светового пятна выбирают равным шагу дискретизации с тем, чтобы целиком заполнить все поле изображения. Такая интерполяция

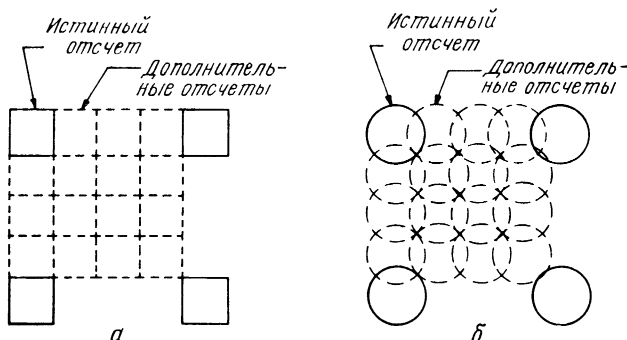


Рис. 4.3.3. Восстановление изображений методом подразвертки.
а — квадратное пятно; *б* — круглое пятно

выполняется просто, но она не является оптимальной. Если в восстанавливающем устройстве удается сформировать очень маленькое световое пятно, то с помощью дополнительной подразвертки можно с некоторой погрешностью синтезировать любую желаемую интерполяционную функцию (рис. 4.3.3). Существуют три метода получения значений промежуточных элементов изображения для восстановления с подразверткой: подгонка пространственной функции, свертка и фильтрация в частотной области. В гл. 11 и 15 подробно рассмотрено практическое применение этих методов.

4.3.2. ИНТЕРПОЛЯЦИОННЫЕ ФУНКЦИИ

На рис. 4.3.4 приведены примеры одномерных интерполяционных функций. Как уже отмечалось, функция sinc (x) обеспечивает точное восстановление, но, как правило, ее трудно сформировать в реальной системе воспроизведения изображений. Простейшей интерполяционной функцией является прямоугольная функция, с помощью которой осуществляется интерполяция отсчетов многочленом нулевого порядка. Треугольная функция обеспечивает линейную интерполяцию первого порядка. Подобную функцию можно рассматривать как свертку двух прямоугольных функций. Свертка треугольной функции с прямоугольной дает колоколооб-

разную интерполяционную функцию, изображенную на рис. 4.3.4, *г*. Повторением этого процесса можно быстро прийти к гауссовой интерполяционной функции, приведенной на рис. 4.3.4, *е*. Многочлены второго и более высокого порядка также пригодны для интерполяции отсчетов. Особенно удобным для интерполяции изображений является кубический *B*-сплайн, поскольку в результате

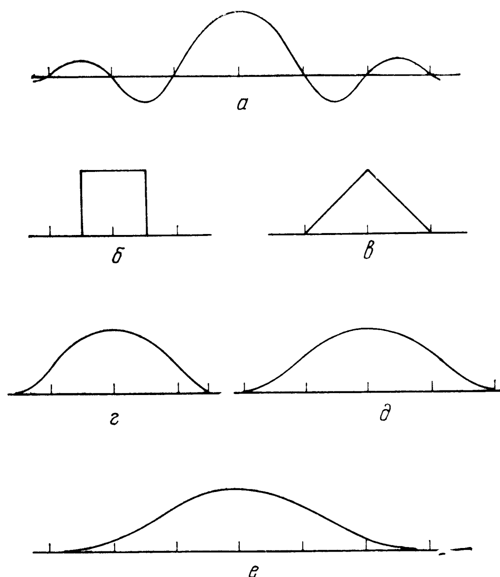


Рис. 4.3.4. Одномерные интерполяционные функции.

а — sinc (*x*); *б* — прямоугольная: *в* — треугольная (свертка двух прямоугольных функций); *г* — колоколообразная (свертка трех прямоугольных функций); *д* — кубический *B*-сплайн (свертка четырех прямоугольных функций); *е* — гауссова

интерполяции получается функция, непрерывная и гладкая в узлах интерполяции. Кубический *B*-сплайн определяется [14] соотношением

$$R(x) = (x)_+^3 - 4(x - \Delta x)_+^3 + 6(x - 2\Delta x)_+^3 - 4(x - 3\Delta x)_+^3, \tag{4.3.1}$$

где

$$(z)_+^m = \begin{cases} z^m, & z > 0, \quad (4.3.2а) \\ 0, & z \leqslant 0. \quad (4.3.2б) \end{cases}$$

Эту функцию, отличную от нуля только на четырех интервалах дискретизации, можно получить, выполнив свертку четырех прямоугольных функций. Рис. 4.3.5 иллюстрирует процесс одномер-

ной интерполяции с использованием функций sinc (x), а также прямоугольных и треугольных функций.

В табл. 4.3.1 даны определения нескольких двумерных разделимых интерполяционных функций, для которых $R(x, y) = R(x)R(y)$. Следует отметить, что операция двумерной линейной интерполяции (или интерполяция многочленами первого порядка), аналогичная операции одномерной линейной интерполяции

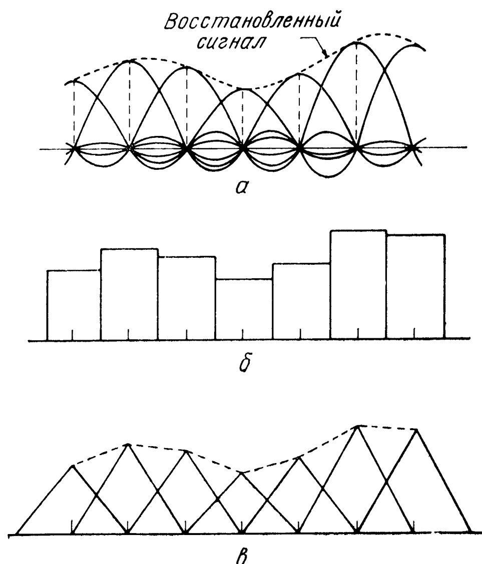


Рис. 4.3.5. Одномерная интерполяция.

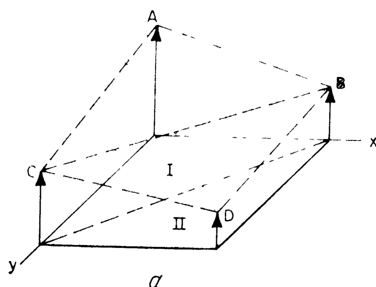
а — функция sinc (x); *б* — прямоугольные функции (интерполяция нулевого порядка); *в* — треугольные функции (интерполяция первого порядка).

(рис. 4.3.5, *в*), отличается от интерполяции с помощью двумерных треугольных функций, представленных в табл. 4.3.1. Эту операцию следует выполнять кусочно-линейным способом, как показано на рис. 4.3.6, *а*. В области *I* отсчеты линейно интерполируются плоскостью, заданной точками A, B и C, тогда как в области *II* они линейно интерполируются плоскостью, заданной точками B, C и D. Непрерывный билинейный способ интерполяции, иллюстрируемый на рис. 4.3.6, *б*, сводится к последовательной линейной интерполяции между парами точек, расположенных на прямых, параллельных осям координат. В результате образуется некоторая поверхность, проходящая через точки A, B, C и D (см. рис. 4.3.6, *б*). Как правило, эта поверхность оказывается неплоской.

Таблица 4.3.1

Двумерные разделимые интерполяционные функции

Функция	Определение
sinc	$R(x, y) = \frac{4}{T_x T_y} \frac{\sin(2\pi x/T_x)}{(2\pi x/T_x)} \frac{\sin(2\pi y/T_y)}{(2\pi y/T_y)}$, $T_x = 2\pi/\omega_{xs}$, $T_y = 2\pi/\omega_{ys}$; $\mathcal{R}(\omega_x, \omega_y) = \begin{cases} 1, \lvert\omega_x\rvert \leqslant \omega_{xs}, & \lvert\omega_y\rvert \leqslant \omega_{ys}, \\ 0 & \text{в противном случае} \end{cases}$
Прямоугольная	$R_0(x, y) = \begin{cases} (1/T_x T_y), \lvert x\rvert \leqslant T_x/2, & \lvert y\rvert \leqslant T_y/2, \\ 0 & \text{в противном случае} \end{cases}$; $\mathcal{R}_0(\omega_x, \omega_y) = \frac{\sin(\omega_x T_x/2)\sin(\omega_y T_y/2)}{(\omega_x T_x/2)(\omega_y T_y/2)}$
Треугольная	$R_1(x, y) = R_0(x, y) * R_0(x, y)$; $\mathcal{R}_1(\omega_x, \omega_y) = \mathcal{R}_0^2(\omega_x, \omega_y)$
Колоколообразная	$R_2(x, y) = R_0(x, y) * R_1(x, y)$; $\mathcal{R}_2(\omega_x, \omega_y) = \mathcal{R}_0^3(\omega_x, \omega_y)$
Кубический B-сплайн	$R_3(x, y) = R_0(x, y) * R_2(x, y)$; $\mathcal{R}_3(\omega_x, \omega_y) = \mathcal{R}_0^4(\omega_x, \omega_y)$
Гауссова	$R(x, y) = [2\pi\sigma_w^2]^{-1} \exp\left\{-\frac{x^2 + y^2}{2\sigma_w^2}\right\}$, $\mathcal{R}(\omega_x, \omega_y) = \exp\left\{-\frac{\sigma_w^2(\omega_x^2 + \omega_y^2)}{2}\right\}$



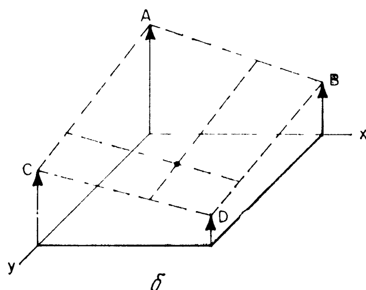


Рис. 4.3.6. Двумерная линейная интерполяция.
а — кусочно-линейная интерполяция; *б* — билинейная интерполяция

4.3.3. ВЛИЯНИЕ НЕИДЕАЛЬНОСТИ ВОССТАНАВЛИВАЮЩИХ ФИЛЬТРОВ

Проанализируем теперь работу реальных систем восстановления непрерывных изображений. Будем предполагать, что входной сигнал такой системы состоит из отсчетов некоторого исходного изображения, полученных на найквистовской частоте с помощью конечного набора дельта-функций. Согласно равенству (4.1.9), восстановленное изображение описывается функцией

$$F_R(x, y) = \sum_{j_1=-J_1}^{J_1} \sum_{j_2=-J_2}^{J_2} F_I(j_1 \Delta x, j_2 \Delta y) R(x - j_1 \Delta x, \quad y - j_2 \Delta y), \tag{4.3.3}$$

где $R(x, y)$ — двумерная интерполяционная функция. В идеальном случае восстановленное изображение должно быть точной копией исходного, т. е.

$$\hat{F}_R(x, y) = \sum_{j=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} F_I(j_1 \Delta x, j_2 \Delta y) R_I(x - j_1 \Delta x, y - j_2 \Delta y), \tag{4.3.4}$$

где $R_I(x, y)$ обозначает оптимальную интерполяционную функцию, например заданную формулой (4.1.14) или (4.1.16). В этом случае ошибка восстановления для произвольной точки изображения равна

$$\mathscr{E}_D(x, y) = \sum_{j_1=-J_1}^{J_1} \sum_{j_2=-J_2}^{J_2} F_I(j_1 \Delta x, j_2 \Delta y) [R(x - j_1 \Delta x, y - j_2 \Delta y) - R_I(x - j_1 \Delta x, y - j_2 \Delta y)]. \tag{4.3.5}$$

Величина ошибки восстановления зависит от двух факторов: а) в реальной системе интерполяционная функция $R(x, y)$ может отличаться от идеальной интерполяционной функции $R_I(x, y)$, б) восстановленное изображение имеет конечные размеры и поэтому интерполяционные функции отсекаются на его границах. В большинстве систем, оперирующих с дискретизованными изображениями, ошибками восстановления вблизи границ обычно пренебрегают, так как они становятся весьма малыми при удалении от границ изображения на несколько шагов дискретизации. Неидеальность интерполяционных функций в принципе приводит к ухудшению резкости изображения, а также к появлению паразитных высокочастотных составляющих.

Влияние неидеальности восстанавливающего фильтра удобно исследовать, рассматривая структуру спектра восстановленного изображения, определяемого соотношением (4.1.11):

$$\mathscr{F}_{\mathscr{R}}(\omega_x, \omega_y) = \frac{4\pi^2}{\Delta x \Delta y} \mathscr{R}(\omega_x, \omega_y) \sum_{j=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathscr{F}_I(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}). \tag{4.3.6}$$

В идеальном случае множитель $\mathscr{R}(\omega_x, \omega_y)$ должен выделять составляющие спектра нулевого порядка при $j_1 = 0$, $j_2 = 0$, а все остальные составляющие должны подавляться. В неидеальном фильтре могут ослабляться составляющие спектра нулевого порядка, что приводит к ухудшению резкости, и, кроме того, гармоники спектров более высокого порядка могут подавляться не полностью, что искажает восстановленное изображение. На

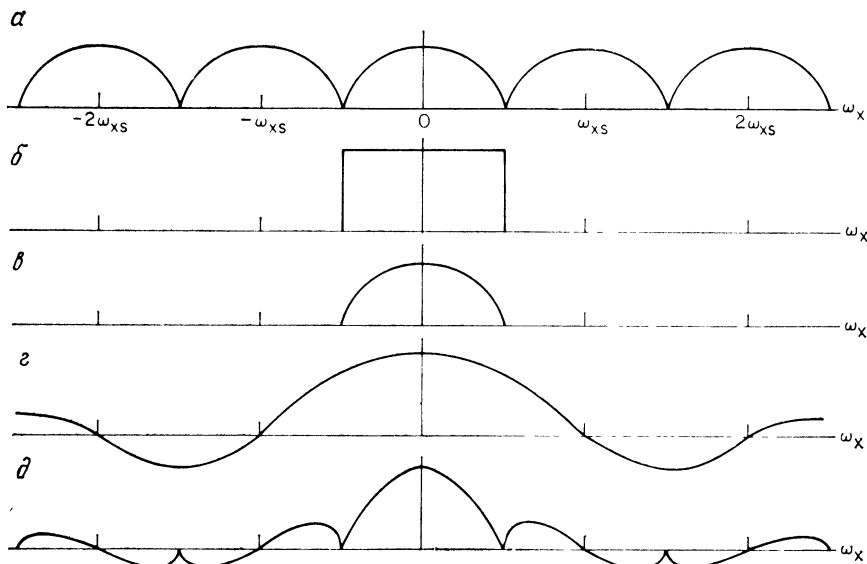


Рис. 4.3.7. Энергетические спектры при идеальном и неидеальном восстановлении.

а — спектр дискретизованного изображения: *б* — частотная характеристика идеального восстанавливающего фильтра; *в* — спектр изображения после идеального восстановления; *г* — частотная характеристика восстанавливающего фильтра с интерполяцией нулевого порядка; *д* — спектр изображения, восстановленного с интерполяцией нулевого порядка.

рис. 4.3.7 приведен пример, показывающий характер искажений спектра изображения при использовании неидеального восстанавливающего фильтра. Типичное сечение спектра дискретизованного изображения представлено на рис. 4.3.7, *а*. При использовании идеального восстанавливающего фильтра, когда интерполяционные функции имеют вид sinc (x), составляющие спектра нулевого порядка выделяются, а все составляющие побочных спектров полностью подавляются (рис. 4.3.7, *в*). На рис. 4.3.7, *г* показана частотная характеристика восстанавливающего фильтра, основанного на интерполяционной функции нулевого порядка, когда значения восстановленных элементов в окрестности точки отсчета устанавливаются равными значению в точке отсчета. В результате

получается спектр (рис. 4.3.7, *д*), искаженный за счет ослабления некоторых составляющих спектра нулевого порядка и появления ложных высокочастотных составляющих.

Повторяя шаги, сделанные при выводе равенства (4.2.21), получим, что потери резкости, вызванные неидеальностью восстанавливающей функции $R(x, y)$, количественно можно выразить с помощью отношения

$$\mathscr{E}_R = (E_{RM} - E_R)/E_{RM}, \tag{4.3.7}$$

где интеграл

$$E_R = \int_{-\omega_{xs}/2}^{\omega_{xs}/2} \int_{-\omega_{ys}/2}^{\omega_{ys}/2} \mathscr{W}_{\mathscr{F}_I}(\omega_x, \omega_y) \, |\mathscr{R}(\omega_x, \omega_y)|^2 \, d\omega_x \, d\omega_y \tag{4.3.8}$$

представляет собой энергию тех составляющих неидеально восстановленного изображения, которые лежат в найквистовском диапазоне, а интеграл

$$E_{RM} = \int_{-\omega_{xs}/2}^{\omega_{xs}/2} \int_{-\omega_{ys}/2}^{\omega_{ys}/2} \mathscr{W}_{\mathscr{F}_I}(\omega_x, \omega_y) \, d\omega_x \, d\omega_y \tag{4.3.9}$$

описывает энергию таких составляющих идеально восстановленного изображения.

Ошибку восстановления, связанную с внесением ложных высокочастотных составляющих спектра, можно определить отношением

$$\mathscr{E}_H = E_H/E_T, \tag{4.3.10}$$

где

$$E_T = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathscr{W}_{\mathscr{F}_P}(\omega_x, \omega_y) \, |\mathscr{R}(\omega_x, \omega_y)|^2 \, d\omega_x \, d\omega_y \tag{4.3.11}$$

есть полная энергия восстановленного изображения, а

$$E_H = E_T - E_R \tag{4.3.12}$$

представляет собой энергию всех спектральных составляющих восстановленного изображения, лежащих вне найквистовского диапазона.

В табл. 4.3.2 приведены значения ошибки восстановления $\mathscr{E}_H$ и коэффициента потери резкости $\mathscr{E}_R$, полученные при восстановлении с использованием некоторых двумерных разделимых интерполяционных функций.

Таблица 4.3.2

Ошибка восстановления и коэффициент потери резкости для различных разделимых двумерных интерполяционных функций

Функция	Коэффициент потери резкости $\mathscr{E}_R$, %	Ошибка восстановления $\mathscr{E}_H$, %
sinc (x)	0,0	0,0
Прямоугольная	26,9	15,7
Треугольная	44,0	3,7
Колоколообразная	55,4	1,1
Кубический B-сплайн	63,2	0,3
Гауссова, $\sigma_w = 3T/8$	38,6	10,3
Гауссова, $\sigma_w = T/2$	54,6	2,0
Гауссова, $\sigma_w = 5T/8$	66,7	0,3

Здесь предполагалось, что спектральная плотность исходного изображения описывается формулами

$$\begin{aligned} \mathscr{W}_{\mathscr{F}_I}(\omega) &= \sqrt{(\omega_s/2)^2 - \omega^2}, \quad \omega^2 \leqslant (\omega_s/2)^2, \\ \mathscr{W}_{\mathscr{F}_I}(\omega) &= 0, \quad \omega^2 > (\omega_s/2)^2. \end{aligned} \tag{4.3.13}$$

Кроме того, предполагалось, что вклад, вносимый в ошибку восстановления побочных спектров при $j_1, j_2 > 2$, пренебрежимо мал. Из таблицы видно, что при восстановлении с использованием прямоугольных интерполяционных функций (функций нулевого порядка) коэффициент потери резкости и ошибка восстановления получаются довольно большими. Применяя интерполяционные функции более высокого порядка, ошибку восстановления можно значительно уменьшить, но за счет дальнейшего ухудшения коэффициента потери резкости.

ЛИТЕРАТУРА

1. Whittaker E. T., On the Functions which are Represented by the Expansions of the Interpolation Theory, *Proc. Roy Soc. Edinburgh, Section A*, **35**, 181—194 (1915).
2. Shannon C. E., Communication in the Presence of Noise, *Proc. IRE*, **37**, 1, 10—21 (January 1949). [Имеется перевод: Шеннон К., Работы по теории информации и кибернетике. — М.: ИЛ, 1963, с. 433—461.]
3. Landau H. J., Sampling, Data Transmission, and the Nyquist Rate, *Proc. IEEE*, **55**, 10, 1701—1706 (October 1967). [Имеется перевод: Ландау, Метод выборок, передача информации и частота Найквиста. — ТИИЭР, 1967, т. 55, № 10.]
4. Goodman J. W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Гудмен Дж., Введение в фурье-оптику. — М.: Мир, 1970.]

Papoulis A., Systems and Transforms with Applications in Optics, McGraw-Hill, New York, 1966. [Имеется перевод: Папулис А., Теория систем и преобразований в оптике. — М.: Мир, 1971.]
Richards P. J., Sampling Positive Functions, *Proc. IEEE* (*Letters*), **54**, 1, 81—82 (January 1966). [Имеется перевод: Ричардс, Дискретное представление положительных функций. — ТИИЭР, 1966, т. 54, № 1.]
Lloyd S. P., A Sampling Theorem for Stationary (Wide Sense) Stochastic Processes, *Trans. Am. Math. Soc.*, **92**, 1, 1—12 (July 1959).
Shapiro H. S., Silverman R. A., Alias-Free Sampling of Random Noise, *J. SIAM*, **8**, 2, 225—248 (June 1960).
Brown J. L., Jr., Bounds for Truncation Error in Sampling Expansions of Band-Limited Signals, *IEEE Trans. Inf. Theory*, **IT-15**, 4, 440—444 (July 1969).
Helms H. D., Thomas J. B., Truncation Error of Sampling Theory Expansions, *Proc. IRE*, **50**, 2, 179—184 (February 1962).
Downing J. J., Data Sampling and Pulse Amplitude Modulation, in: Aerospace Telemetry, Stiltz H. L., Ed., Prentice-Hall, Englewood Cliffs, N. J., 1961.
Childers D. G., Study and Experimental Investigation on Sampling Rate and Aliasing in Time Division Telemetry Systems, *IRE Trans. Space Electronics and Telemetry*, **SET-8**, 267—283 (December 1962).
O'Neill E. L., Introduction to Statistical Optics, Addison-Wesley, Reading, Mass., 1963. [Имеется перевод: О'Нейл Э., Введение в статистическую оптику. — М.: Мир, 1966.]
Greville T. N. E., Introduction to Spline Functions, in: Theory and Applications of Spline Functions (Greville T. N. E., Ed.), Academic Press, New York, 1969.

Глава 5

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ДИСКРЕТНЫХ ИЗОБРАЖЕНИЙ

В гл. 2 были рассмотрены вопросы, связанные с математическим описанием непрерывных изображений. В настоящей главе даны способы формального представления дискретных изображений с использованием как детерминированных, так и статистических моделей.

5.1. ДЕЙСТВИЯ С ВЕКТОРАМИ И МАТРИЦАМИ

В данном разделе коротко рассмотрены встречающиеся в тексте математические действия, выполняемые с векторами и матрицами. Строгий вывод и доказательства теорем и положений, приведенных ниже, можно найти в литературе [1—5].

Вектор

Вектор-столбец $\mathbf{f}$ размера $N \times 1$ представляет собой совокупность элементов $f(n)$, где $n = 1, 2, \ldots, N$, расположенных в виде вертикального столбца

$$\mathbf{f} = \begin{bmatrix} f(1) \\ f(2) \\ \cdots \\ f(j) \\ \cdots \\ f(N) \end{bmatrix}. \tag{5.1.1}$$

Вектор-строка $\mathbf{h}$ размера $1 \times N$ представляет собой упорядоченную совокупность элементов $h(n)$, где $n = 1, 2, \ldots N$, расположенных в виде горизонтальной строки

$$\mathbf{h} = [h(1), h(2), \ldots, h(j), \ldots, h(N)]. \tag{5.1.2}$$

В книге полужирными строчными буквами будут, как правило, обозначаться вектор-столбцы. Вектор-строка будет обозначаться как транспонированный вектор-столбец:

$$\mathbf{f}^T = [f(1), f(2), \ldots, f(j), \ldots, f(N)]. \tag{5.1.3}$$

Матрица

Матрица $\mathbf{F}$ размера $M \times N$ представляет собой совокупность элементов $F(m, n)$, где $m = 1, 2, \ldots, M$ и $n = 1, 2, \ldots, N$, расположенных в виде строк и столбцов двумерной таблицы

$$\mathbf{F} = \begin{bmatrix} F(1, 1) & F(1, 2) & \ldots & F(1, N) \\ F(2, 1) & F(2, 2) & \ldots & F(2, N) \\ \cdot & \cdot & \cdot & \cdot \\ F(M, 1) & F(M, 2) & \ldots & F(M, N) \end{bmatrix}. \qquad (5.1.4)$$

Символ $\mathbf{0}$ обозначает нулевую матрицу, все элементы которой равны нулю. Диагональная матрица — это квадратная матрица (когда $M = N$), все элементы которой, не лежащие на главной диагонали, равны нулю, т. е. $F(m, n) = 0$, если $m \neq n$. Единичная матрица, обозначаемая символом $\mathbf{I}$, есть диагональная матрица, все диагональные элементы которой равны единице. Индекс при символе единичной матрицы указывает ее размеры; $\mathbf{I}_N$ обозначает единичную матрицу размера $N \times N$. Матрица $\mathbf{F}$ может быть разделена на блоки (подматрицы) $\mathbf{F}_{pq}$:

$$\mathbf{F} = \left[\begin{array}{c:c:c:c} \mathbf{F}_{1,1} & \mathbf{F}_{1,2} & \ldots & \mathbf{F}_{1,Q} \\ \hdashline \ldots & \ldots & \ldots & \ldots \\ \hdashline \mathbf{F}_{P,1} & \mathbf{F}_{P,2} & \ldots & \mathbf{F}_{P,Q} \end{array}\right]. \qquad (5.1.5)$$

Сложение матриц

Сумма двух матриц $\mathbf{C} = \mathbf{A} + \mathbf{B}$ определена только в том случае, когда обе матрицы имеют одинаковые размеры. Матрица $\mathbf{C}$ — сумма матриц $\mathbf{A}$ и $\mathbf{B}$, имеет размеры $M \times N$, а ее элементы $C(m, n) = A(m, n) + B(m, n)$.

Умножение матриц

Произведение двух матриц $\mathbf{C} = \mathbf{AB}$ определено только тогда, когда число столбцов матрицы $\mathbf{A}$ равно числу строк матрицы $\mathbf{B}$. При умножении матрицы $\mathbf{A}$ размера $M \times P$ на матрицу $\mathbf{B}$ размера $P \times N$ получается матрица $\mathbf{C}$ размера $M \times N$, элементы которой определяются равенством

$$C(m, n) = \sum_{p=1}^{P} A(m, p) B(p, n). \qquad (5.1.6)$$

При умножении матрицы $\mathbf{A}$ на скаляр k получается матрица $\mathbf{C} = k\mathbf{A}$, элементы которой $C(m, n) = kA(m, n)$.

Обращение матриц

Если $\mathbf{A}$ — квадратная матрица, то матрица, обратная относительно нее и обозначаемая как $\mathbf{A}^{-1}$, обладает следующими свойствами: $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ и $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Если матрица $\mathbf{A}^{-1}$ существует, то матрица $\mathbf{A}$ называется неособенной (невырожденной). В противном случае она называется особенной (вырожденной). Если у некоторой матрицы есть обратная, то эта обратная матрица единственна. Матрица, обратная относительной обратной, совпадает с исходной матрицей, т. е.

$$[\mathbf{A}^{-1}]^{-1} = \mathbf{A}. \tag{5.1.7}$$

Если матрицы $\mathbf{A}$ и $\mathbf{B}$ неособенные, то

$$[\mathbf{A}\mathbf{B}]^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{5.1.8}$$

Если матрица $\mathbf{A}$ неособенная, а скаляр $k \neq 0$, то

$$[k\mathbf{A}]^{-1} = (1/k)\,\mathbf{A}^{-1}. \tag{5.1.9}$$

Обращение особенных квадратных матриц и неквадратных матриц будет рассмотрено в гл. 8. Матрицу, обратную относительно блочной квадратной матрицы

$$\mathbf{F} = \left[\begin{array}{c|c} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \hline \mathbf{F}_{21} & \mathbf{F}_{22} \end{array}\right], \tag{5.1.10}$$

можно представить в виде

$$\mathbf{F}^{-1} = \left[\begin{array}{c|c} [\mathbf{F}_{11} - \mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{21}]^{-1} & -\mathbf{F}_{11}^{-1}\mathbf{F}_{12}\,[\mathbf{F}_{22} - \mathbf{F}_{21}\mathbf{F}_{11}^{-1}\mathbf{F}_{12}]^{-1} \\ \hline -\mathbf{F}_{22}^{-1}\mathbf{F}_{21}\,[\mathbf{F}_{11} - \mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{21}]^{-1} & [\mathbf{F}_{22} - \mathbf{F}_{21}\mathbf{F}_{11}^{-1}\mathbf{F}_{11}]^{-1} \end{array}\right] \tag{5.1.11}$$

при условии, что матрицы $\mathbf{F}_{11}$ и $\mathbf{F}_{22}$ не являются особенными.

Транспонирование матриц

При транспонировании матрицы $\mathbf{A}$ размера $M \times N$ образуется матрица размера $N \times M$, которую обозначают через $\mathbf{A}^T$. Строки матрицы $\mathbf{A}^T$ совпадают со столбцами, а столбцы — со строками матрицы $\mathbf{A}$. Для любой матрицы $\mathbf{A}$

$$[\mathbf{A}^T]^T = \mathbf{A}. \tag{5.1.12}$$

Если $\mathbf{A} = \mathbf{A}^T$, то матрицу $\mathbf{A}$ называют симметричной. Для любых матриц $\mathbf{A}$ и $\mathbf{B}$

$$[\mathbf{A}\mathbf{B}]^T = \mathbf{B}^T\mathbf{A}^T. \tag{5.1.13}$$

Если матрица **A** неособенная, то матрица $\mathbf{A}^T$ также неособенная и

$$[\mathbf{A}^T]^{-1} = [\mathbf{A}^{-1}]^T. \qquad (5.1.14)$$

Прямое произведение матриц

Левое прямое произведение матрицы **A** размера $P \times Q$ на матрицу **B** размера $M \times N$ представляет собой матрицу размера $PM \times QN$:

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} B(1,\ 1)\mathbf{A} & B(1,\ 2)\mathbf{A} & \dots & B(1,\ N)\mathbf{A} \\ B(2,\ 1)\mathbf{A} & B(2,\ 2)\mathbf{A} & \dots & B(2,\ N)\mathbf{A} \\ \cdots & \cdots & \cdots & \cdots \\ B(M,\ 1)\mathbf{A} & B(M,\ 2)\mathbf{A} & \dots & B(M,N)\mathbf{A} \end{bmatrix}. \qquad (5.1.15)$$

Аналогично можно определить правое прямое произведение. В этой книге будет использоваться только левое прямое произведение. Прямые произведения $\mathbf{B} \otimes \mathbf{A}$ и $\mathbf{A} \otimes \mathbf{B}$ могут различаться между собой. Ниже указаны свойства операций умножения, сложения, транспонирования и обращения прямого произведения матриц:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}), \qquad (5.1.16)$$

$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}, \qquad (5.1.17)$$

$$[\mathbf{A} \otimes \mathbf{B}]^T = \mathbf{A}^T \otimes \mathbf{B}^T, \qquad (5.1.18)$$

$$[\mathbf{A} \otimes \mathbf{B}]^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}. \qquad (5.1.19)$$

След матрицы

След квадратной матрицы **F** размера $N \times N$ равен сумме ее диагональных элементов и обозначается как

$$\mathrm{tr}[\mathbf{F}] = \sum_{n=1}^{N} F(n,\ n). \qquad (5.1.20)$$

Если **A** и **B** — квадратные матрицы, то

$$\mathrm{tr}[\mathbf{AB}] = \mathrm{tr}[\mathbf{BA}]. \qquad (5.1.21)$$

След прямого произведения двух матриц равен

$$\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A})\,\mathrm{tr}(\mathbf{B}). \qquad (5.1.22)$$

Норма вектора

Евклидовой нормой вектора **f** размера $N \times 1$ называется скаляр, определяемый как

$$\|\mathbf{f}\| = \mathbf{f}^T\mathbf{f}. \qquad (5.1.23)$$

Норма матрицы

Евклидовой нормой матрицы **F** размера $M \times N$ называется скаляр, определяемый следующим образом:

$$\| \mathbf{F} \| = \operatorname{tr} [\mathbf{F}^T \mathbf{F}]. \qquad (5.1.24)$$

Ранг матрицы

Матрица **A** размера $N \times N$ имеет ранг R, если наибольший из всех ее квадратных неособенных блоков имеет размер $R \times R$. Понятие о ранге используется при обращении матриц. Если матрицы **A** и **B** неособенные, а **C** — произвольная матрица, то

$$\text{Ранг } [\mathbf{C}] = \text{Ранг } [\mathbf{AC}] = \text{Ранг } [\mathbf{CB}] = \text{Ранг } [\mathbf{ACB}]. \qquad (5.1.25)$$

Ранг произведения матриц **A** и **B** удовлетворяет неравенствам

$$\text{Ранг } [\mathbf{AB}] \leqslant \text{Ранг } [\mathbf{A}], \qquad (5.1.26а)$$

$$\text{Ранг } [\mathbf{AB}] \leqslant \text{Ранг } [\mathbf{B}]. \qquad (5.1.26б)$$

Ранг суммы матриц **A** и **B** удовлетворяет неравенству

$$\text{Ранг } [\mathbf{A} + \mathbf{B}] \leqslant \text{Ранг } [\mathbf{A}] + \text{Ранг } [\mathbf{B}]. \qquad (5.1.27)$$

Скалярное произведение векторов

Скалярным произведением векторов **f** и **g** размера $N \times 1$ является скаляр

$$k = \mathbf{g}^T \mathbf{f} = \mathbf{f}^T \mathbf{g}, \qquad (5.1.28)$$

или

$$k = \sum_{n=1}^{N} g(n) f(n). \qquad (5.1.29)$$

Матричное произведение векторов

Матричным произведением вектора g размера $M \times 1$ на вектор **f** размера $N \times 1$ является матрица

$$\mathbf{A} = \mathbf{g}\mathbf{f}^T, \qquad (5.1.30)$$

где $\mathbf{A}(m, n) = g(m) f(n)$.

Квадратичная форма

Квадратичной формой вектора **f** размера $N \times 1$ является скаляр

$$k = \mathbf{f}^T \mathbf{A} \mathbf{f}, \qquad (5.1.31)$$

где **A** — матрица размера $N \times N$. Часто матрицу **A** берут симметричной.

Векторная производная

Производная от скалярного произведения $\mathbf{a}^T\mathbf{x}$ по $\mathbf{x}$ есть

$$\partial[\mathbf{a}^T\mathbf{x}]/\partial\mathbf{x} = \mathbf{a}, \qquad (5.1.32)$$

а производная от скалярного произведения $\mathbf{x}^T\mathbf{a}$ по вектору $\mathbf{x}$ равна

$$\partial[\mathbf{x}^T\mathbf{a}]/\partial\mathbf{x} = \mathbf{a}. \qquad (5.1.33)$$

Производная от квадратичной формы $\mathbf{x}^T\mathbf{A}\mathbf{x}$ по $\mathbf{x}$ есть [1])

$$\partial[\mathbf{x}^T\mathbf{A}\mathbf{x}]/\partial\mathbf{x} = 2\mathbf{A}\mathbf{x}. \qquad (5.1.34)$$

5.2. СИНГУЛЯРНОЕ РАЗЛОЖЕНИЕ МАТРИЦЫ

Известно, что любую матрицу $\mathbf{F}$ размера $M \times N$, имеющую ранг R, можно представить в виде взвешенной суммы матриц единичного ранга размера $M \times N$. Такое представление называется сингулярным разложением [6—8]. В последующих разделах будет рассмотрено применение этого метода для обработки изображений.

При сингулярном разложении используют унитарную матрицу $\mathbf{U}$ размера $M \times N$ и унитарную матрицу $\mathbf{V}$ размера $N \times N$, такие, что

$$\mathbf{U}^T\mathbf{F}\mathbf{V} = \mathbf{\Lambda}^{1/2}, \qquad (5.2.1)$$

где матрица

$$\mathbf{\Lambda}^{1/2} = \begin{array}{c} \begin{array}{cc} \overbrace{\hphantom{\lambda^{1/2}(1)\quad\ddots\quad\lambda^{1/2}(R)}}^{R} & \overbrace{\hphantom{\mathbf{0}\;\;}}^{N-R} \end{array} \\ \left[\begin{array}{ccc|c} \lambda^{1/2}(1) & & & \\ & \ddots & & \mathbf{0} \\ & & \lambda^{1/2}(R) & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right\} R \\ \}\ M-R \end{array} \qquad (5.2.2)$$

имеет размеры $M \times N$, а ее диагональные элементы $\lambda^{1/2}(j)$ называются сингулярными значениями матрицы $\mathbf{F}$. Поскольку матрицы $\mathbf{U}$ и $\mathbf{V}$ унитарны, то $\mathbf{U}\mathbf{U}^T = \mathbf{I}_M$ и $\mathbf{V}\mathbf{V}^T = \mathbf{I}_N$. Поэтому

$$\mathbf{F} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^T. \qquad (5.2.3)$$

Столбцы унитарной матрицы $\mathbf{U}$ являются собственными векторами $\mathbf{u}_m$ симметричной матрицы $\mathbf{F}\mathbf{F}^T$, т. е.

$$\mathbf{U}^T[\mathbf{F}\mathbf{F}^T]\mathbf{U} = \begin{array}{c} \begin{array}{cc} \overbrace{\hphantom{\lambda(1)\quad\ddots\quad\lambda(R)}}^{R} & \overbrace{\hphantom{\mathbf{0}\;\;}}^{M-R} \end{array} \\ \left[\begin{array}{ccc|c} \lambda(1) & & & \\ & \ddots & & \mathbf{0} \\ & & \lambda(R) & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right\} R \\ \}\ M-R \end{array}, \qquad (5.2.4)$$

---

[1]) Формула (5.1.34) справедлива в случае симметричной матрицы $\mathbf{A}$. — *Прим. ред.*

где $\lambda(j)$ — ненулевые собственные значения матрицы $\mathbf{F}\mathbf{F}^T$. Аналогично строки матрицы $\mathbf{V}$ являются собственными векторами $\mathbf{v}_n$ симметричной матрицы $\mathbf{F}^T\mathbf{F}$, т. е.

$$\mathbf{V}^T[\mathbf{F}^T\mathbf{F}]\mathbf{V} = \begin{array}{c} \overbrace{\hphantom{\lambda(1)\quad\ddots\quad\lambda(R)}}^{R}\;\overbrace{\hphantom{\quad\mathbf{0}\quad}}^{N-R} \\ \left[\begin{array}{ccc|c} \lambda(1) & & & \\ & \ddots & & \mathbf{0} \\ & & \lambda(R) & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right\} R \\ \} N-R \end{array}, \quad (5.2.5)$$

где $\lambda(j)$ — соответствующие ненулевые собственные значения матрицы $\mathbf{F}^T\mathbf{F}$. Нетрудно проверить, что равенство (5.2.3) согласуется с (5.2.4) и (5.2.5).

Разложение матрицы $\mathbf{F}$, задаваемое соотношением (5.2.3), можно представить в виде ряда

$$\mathbf{F} = \sum_{j=1}^{R} \lambda^{1/2}(j)\, \mathbf{u}_j \mathbf{v}_j^T. \quad (5.2.6)$$

Матричные произведения собственных векторов $\mathbf{u}_j\mathbf{v}_j^T$ образуют набор матриц единичного ранга, каждая из которых умножается на весовой множитель, являющийся соответствующим сингулярным значением матрицы $\mathbf{F}$. Согласованность разложения (5.2.6) с вышеприведенными соотношениями можно показать, подставив его в равенство (5.2.1). В результате получается

$$\mathbf{\Lambda}^{1/2} = \mathbf{U}^T\mathbf{F}\mathbf{V} = \sum_{j=1}^{R} \lambda^{1/2}(j)\, \mathbf{U}^T\mathbf{u}_j\mathbf{v}_j^T\mathbf{V}. \quad (5.2.7)$$

Заметим, что произведение $\mathbf{U}^T\mathbf{u}_j$ дает вектор-столбец, j-й элемент которого равен единице, а все остальные — нули. Вектор-строка, получающаяся в результате вычисления произведения $\mathbf{v}_j^T\mathbf{V}$, имеет аналогичный вид. Поэтому в правой части равенства (5.2.7) образуется диагональная матрица, элементы которой равны сингулярным значениям матрицы $\mathbf{F}$.

Матричное разложение (5.2.3) и эквивалентное представление в виде ряда (5.2.6) можно найти для любой матрицы. Поэтому такое разложение можно непосредственно применить для обработки дискретных изображений, представленных в виде матриц. Кроме того, этими формулами можно воспользоваться для разложения матриц линейных преобразований изображений. Применение метода сингулярного разложения для исправления и кодирования изображений рассмотрено в последующих главах книги.

5.3. ВЕКТОРНОЕ ПРЕДСТАВЛЕНИЕ ИЗОБРАЖЕНИЙ

В гл. 1 яркость, координата цвета или какой-то другой подходящий параметр, описывающий изображение, представлялись непрерывной функцией $F(x, y, t)$. С помощью методов дискрети-

зации изображений, рассмотренных в гл. 4, непрерывное изображение, зафиксированное в некоторый момент времени, можно представить в виде массива отсчетов $F(j_1, j_2)$ в некоторой прямоугольной области $-J_i \leqslant j_i \leqslant J_i$. Часто этот массив полезно рассматривать как матрицу с $N_1 \times N_2$ элементами:

$$\mathbf{F} = [F(n_1, n_2)], \tag{5.3.1}$$

где $1 \leqslant n_i \leqslant N_i$, а индексы отсчетов перенумерованы так, как это принято в теории матриц.

Для облегчения анализа удобно перейти от матричного представления изображения к векторному, собирая элементы столбцов (или строк) матрицы $\mathbf{F}$ в один длинный вектор [9]. Формально эту операцию можно представить с помощью вспомогательного вектора $\mathbf{v}_n$ размера $N_2 \times 1$ и матрицы $\mathbf{N}_n$ размера $N_1N_2 \times N_1$, определяемых следующим образом:

$$\mathbf{v}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} 1 \\ \vdots \\ n-1 \\ n \\ n+1 \\ \vdots \\ N_2 \end{matrix}, \qquad \mathbf{N}_n = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{I} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \begin{matrix} 1 \\ \vdots \\ n-1 \\ n \\ n+1 \\ \vdots \\ N_2 \end{matrix} \tag{5.3.2}$$

В этом случае матрица $\mathbf{F}$ будет представлена в векторной форме с помощью операции упорядочения

$$\mathbf{f} = \sum_{n=1}^{N_2} \mathbf{N}_n \mathbf{F} \mathbf{v}_n. \tag{5.3.3}$$

Вектор $\mathbf{v}_n$ выделяет n-й столбец матрицы $\mathbf{F}$, а матрица $\mathbf{N}_n$ помещает этот столбец на место, отведенное для n-го отрезка вектора $\mathbf{f}$. Таким образом, вектор $\mathbf{f}$ содержит все элементы матрицы $\mathbf{F}$, последовательно считанные по столбцам. Обратная операция преобразования вектора $\mathbf{f}$ в матрицу $\mathbf{F}$ описывается соотношением

$$\mathbf{F} = \sum_{n=1}^{N_2} \mathbf{N}_n^T \mathbf{f} \mathbf{v}_n^T. \tag{5.3.4}$$

С помощью формул (5.3.3) и (5.3.4) легко установить связь между матричным и векторным представлениями двумерного массива. Достоинствами представления изображения в векторной форме являются большая компактность обозначений и возможность непосредственного использования методов, разработанных для обработки одномерных сигналов. Следует отметить, что выражения (5.3.3) и (5.3.4) не только описывают лексикографическую связь

между матрицей и вектором, но и определяют некоторые операторы, которыми можно пользоваться при математическом анализе. В последующих разделах приводятся многочисленные примеры применения этих операторов упорядочения.

5.4. СТАТИСТИЧЕСКОЕ ОПИСАНИЕ ДИСКРЕТНЫХ ИЗОБРАЖЕНИЙ [5, 10]

Статистические методы описания непрерывных изображений, приведенные в гл. 1, можно непосредственно применить и для описания дискретных изображений. В данном разделе получены выражения для моментов дискретных изображений. Модели совместных плотностей вероятностей приведены в следующем разделе.

Среднее значение матрицы, описывающей дискретное изображение, представляет собой матрицу

$$E\{\mathbf{F}\} = [E\{F(n_1, n_2)\}]. \tag{5.4.1}$$

Если эта матрица разверткой по столбцам преобразована в вектор, то среднее значение этого вектора есть

$$\eta_f = E\{\mathbf{f}\} = \sum_{n=1}^{N_2} \mathbf{N}_n E\{\mathbf{F}\} \mathbf{v}_n. \tag{5.4.2}$$

Корреляция двух элементов изображения с координатами (n_1, n_2) и (n_3, n_4) определяется как

$$R(n_1, n_2; n_3, n_4) = E\{F(n_1, n_2) F^*(n_3, n_4)\}. \tag{5.4.3}$$

Ковариация двух элементов изображения есть

$$K(n_1, n_2; n_3, n_4) = E\{[F(n_1, n_2) - E\{F(n_1, n_2)\}] \times \\ \times [F^*(n_3, n_4) - E\{F^*(n_3, n_4)\}]\}. \tag{5.4.4}$$

И наконец, дисперсия элемента изображения равна

$$\sigma^2(n_1, n_2) = K(n_1, n_2; n_1, n_2). \tag{5.4.5}$$

Если матрица изображения преобразована в вектор $\mathbf{f}$, то корреляционную матрицу этого вектора можно выразить через корреляции элементов матрицы $\mathbf{F}$:

$$\mathbf{R}_f = E\{\mathbf{f}\mathbf{f}^{*T}\} = E\left\{\left[\sum_{m=1}^{N_2} \mathbf{N}_m \mathbf{F} \mathbf{v}_m\right]\left[\sum_{n=1}^{N_2} \mathbf{v}_n^T \mathbf{F}^{*T} \mathbf{N}_n^T\right]\right\}, \tag{5.4.6а}$$

или

$$\mathbf{R}_f = \sum_{m=1}^{N_2} \sum_{n=1}^{N_2} \mathbf{N}_m E\{\mathbf{F} \mathbf{v}_m \mathbf{v}_n^T \mathbf{F}^{*T}\} \mathbf{N}_n^T. \tag{5.4.6б}$$

Выражение

$$E\{\mathbf{F}\mathbf{v}_m\mathbf{v}_n^T\mathbf{F}^{*T}\} = \mathbf{R}_{m,\,n} \tag{5.4.7}$$

представляет собой корреляционную матрицу m-го и n-го столбцов матрицы $\mathbf{F}$ и имеет размеры $N_1 \times N_1$. Следовательно, $\mathbf{R}_f$ можно представить в виде блочной матрицы

$$\mathbf{R}_f = \begin{bmatrix} \mathbf{R}_{1,\,1} & \mathbf{R}_{1,\,2} & \dots & \mathbf{R}_{1,\,N_2} \\ \mathbf{R}_{2,\,1} & \mathbf{R}_{2,\,2} & \dots & \mathbf{R}_{2,\,N_2} \\ \dots & \dots & \dots & \dots \\ \mathbf{R}_{N_2,\,1} & \mathbf{R}_{N_2,\,2} & \dots & \mathbf{R}_{N_2,\,N_2} \end{bmatrix}. \tag{5.4.8}$$

Ковариационную матрицу вектора $\mathbf{f}$ можно получить на основе его корреляционной матрицы и вектора средних значений с помощью соотношения

$$\mathbf{K}_f = \mathbf{R}_f - \boldsymbol{\eta}_f\boldsymbol{\eta}_f^{*T}. \tag{5.4.9}$$

Матрица дисперсий $\mathbf{V}_F$ массива чисел $F(n_1, n_2)$ по определению является матрицей, элементы которой равны дисперсиям соответствующих элементов массива. Элементы матрицы $\mathbf{V}_F$ можно непосредственно выделить из блоков матрицы $\mathbf{K}_f$:

$$\mathbf{V}_F(n_1, n_2) = \mathbf{K}_{n_2 n_2}(n_1, n_1). \tag{5.4.10}$$

Если дискретное изображение представляется массивом, стационарным в широком смысле, то его корреляционную функцию можно записать в виде

$$R(n_1, n_2; n_3, n_4) = R(n_1 - n_3, n_2 - n_4) = R(j, k), \tag{5.4.11}$$

где $j = n_1 - n_3$ и $k = n_2 - n_4$. Соответственно блоки ковариационной матрицы (5.4.9) будут связаны соотношениями

$$\mathbf{K}_{m,\,n} = \mathbf{K}_k, \quad m \geqslant n, \tag{5.4.12а}$$

$$\mathbf{K}_{m,\,n} = \mathbf{K}_k^*, \quad m < n, \tag{5.4.12б}$$

где $k = |m - n| + 1$. Таким образом, для стационарного в широком смысле массива

$$\mathbf{K}_f = \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 & \mathbf{K}_3 & \dots & \mathbf{K}_{N_2} \\ \mathbf{K}_2^* & \mathbf{K}_1 & \mathbf{K}_2 & \dots & \mathbf{K}_{N_2-1} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{K}_{N_2}^* & \mathbf{K}_{N_2-1}^* & \mathbf{K}_{N_2-2}^* & \dots & \mathbf{K}_1 \end{bmatrix}. \tag{5.4.13}$$

Матрица (5.4.13) является блочно-тёплицевой [11]. Наконец, если корреляционную функцию изображения можно записать в виде произведения корреляционных функций строк и столбцов, то ковариационную матрицу вектора $\mathbf{f}$, представляющего

изображение, можно записать в виде прямого произведения ковариационных матриц для строк и столбцов:

$$\mathbf{K}_f = \mathbf{K}_C \otimes \mathbf{K}_R =$$

$$= \begin{bmatrix} \mathbf{K}_R(1,\ 1)\,\mathbf{K}_C & \mathbf{K}_R(1,\ 2)\,\mathbf{K}_C & \dots & \mathbf{K}_R(1,\ N_2)\,\mathbf{K}_C \\ \mathbf{K}_R(2,\ 1)\,\mathbf{K}_C & \mathbf{K}_R(2,\ 2)\,\mathbf{K}_C & \dots & \mathbf{K}_R(2,\ N_2)\,\mathbf{K}_C \\ \dots & \dots & \dots & \dots \\ \mathbf{K}_R(N_2,\ 1)\,\mathbf{K}_C & \mathbf{K}_R(N_2,\ 2)\,\mathbf{K}_C & \dots & \mathbf{K}_R(N_2,\ N_2)\,\mathbf{K}_C \end{bmatrix}, \quad (5.4.14)$$

где $\mathbf{K}_C$ — ковариационная матрица столбцов матрицы **F**, имеющая размеры $N_1 \times N_1$, а $\mathbf{K}_R$ — ковариационная матрица строк матрицы **F** с размерами $N_2 \times N_2$ [1]).

Рассмотрим случай, когда ковариационная матрица строк матрицы **F** имеет следующий вид:

$$\mathbf{K}_R = \sigma^2 \begin{bmatrix} 1 & \rho_R & \rho_R^2 & \rho_R^3 & \dots & \rho_R^{N_2-1} \\ \rho_R & 1 & \rho_R & \rho_R^2 & \dots & \rho_R^{N_2-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_R^{N_2-1} & \dots & \dots & \dots & \dots & 1 \end{bmatrix}, \quad (5.4.15)$$

где σ^2 — дисперсия элементов изображения. Эта ковариационная матрица — аналог непрерывной автоковариационной функции вида $\exp\{-\alpha|x|\}$ — описывает марковский процесс. На рис. 5.4.1 приведены полученные Дэвиссоном [12] значения коэффициентов корреляции элементов строки типичного изображения. Экспериментальные точки хорошо аппроксимируются ковариационной функцией марковского процесса с параметром $\rho = 0{,}953$. Аналогично значения коэффициентов корреляции в направлении, перпендикулярном к строкам, хорошо согласуются с марковской ковариационной функцией при $\rho = 0{,}965$. Если ковариационная функция может быть представлена в виде (5.4.14), то коэффициенты корреляции по диагонали должны быть равны произведению соответствующих коэффициентов корреляции вдоль строк изображения и в направлении, перпендикулярном к ним. В данном примере оказалось, что такая аппроксимация является достаточно точной в области от нуля до пяти шагов дискретизации.

По аналогии с непрерывным энергетическим спектром (1.8.11) можно определить дискретную спектральную плотность дискретного стационарного двумерного случайного поля, представляющего изображение, как результат двумерного дискретного преобра-

---

[1]) Выражение (5.4.14) должно иметь следующий вид:

$$(1/\sigma^2)\,\mathbf{K}_f = [(1/\sigma^2)\,\mathbf{K}_C] \otimes [(1/\sigma^2)\,\mathbf{K}_R],$$

где σ^2 — дисперсия элементов изображения. — *Прим. ред.*

зования Фурье [1]) автокорреляционной функции этого поля. Тогда в силу равенства (5.4.11) будем иметь

$$\mathscr{W}(u, v) = \frac{1}{\sqrt{N_1 N_2}} \sum_{j_1=0}^{N_1-1} \sum_{k=0}^{N_2-1} R(j, k) \exp\left\{-2\pi i \left(\frac{ju}{N_1} + \frac{kv}{N_2}\right)\right\}. \tag{5.4.16}$$

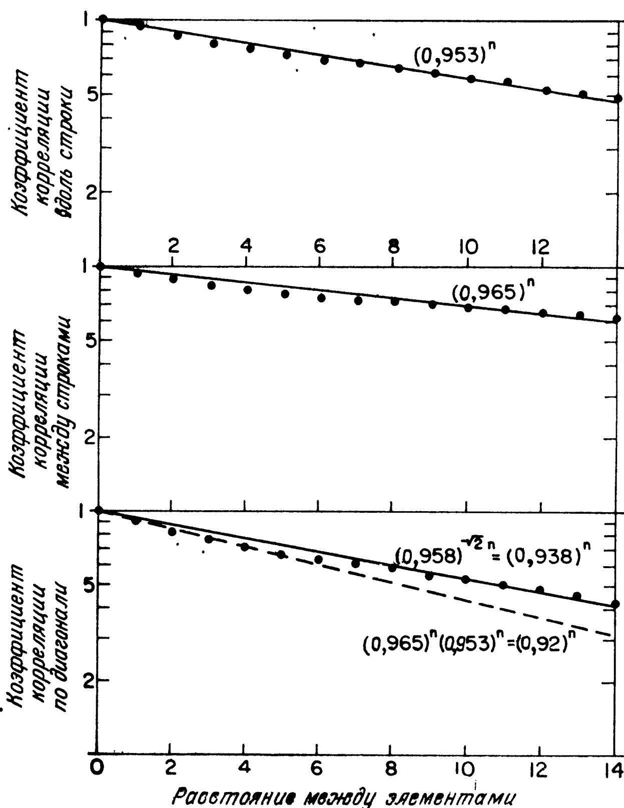


Рис. 5.4.1. Пример корреляционных зависимостей между соседними элементами изображения.

На рис. 5.4.2 приведены энергетические спектры марковских процессов.

---

[1]) Свойства двумерных дискретных преобразований Фурье рассматриваются в гл. 10.

5.5. МОДЕЛИ ПЛОТНОСТЕЙ ВЕРОЯТНОСТИ ДИСКРЕТНЫХ ИЗОБРАЖЕНИЙ

Дискретное изображение, представленное массивом $F(n_1, n_2)$, можно полностью описать с помощью совместной плотности ве-

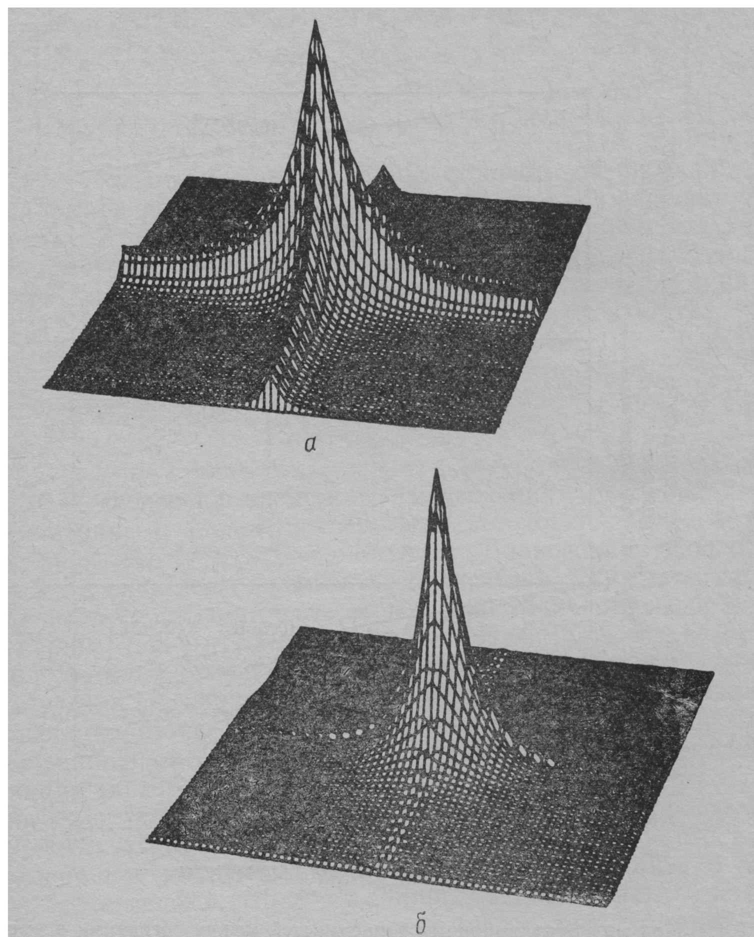


Рис. 5.4.2. Энергетические спектры марковских процессов, моделирующих изображения ($N = 256$, по вертикальной оси логарифмический масштаб).
а — разделимый спектр; *б* — спектр с круговой симметрией

роятности его элементов. При задании $F(n_1, n_2)$ в матричной форме эта плотность записывается как

$$p(\mathbf{F}) \equiv p\{F(1, 1), F(2, 1), \ldots, F(N_1, N_2)\}, \qquad (5.5.1а)$$

а в векторной форме — как

$$p(\mathbf{f}) = p\{f(1),\ f(2),\ \ldots,\ f(Q)\}, \qquad (5.5.1б)$$

где $Q = N_1 \cdot N_2$ определяет порядок совместной плотности. Если все элементы изображения статистически независимы, то совместная плотность вероятности равна произведению одномерных безусловных плотностей

$$p(\mathbf{f}) = p\{f(1)\}\, p\{f(2)\}\ \ldots\ p\{f(Q)\}. \qquad (5.5.2)$$

Наиболее распространенным видом совместной плотности вероятности является гауссова плотность

$$p(\mathbf{f}) = (2\pi)^{-Q/2}\, |\mathbf{K}_f|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{f} - \boldsymbol{\eta}_f)^T\, \mathbf{K}_f^{-1}\, (\mathbf{f} - \boldsymbol{\eta}_f)\right\}, \qquad (5.5.3)$$

где $\mathbf{K}_f$ — ковариационная матрица вектора $\mathbf{f}$, $\boldsymbol{\eta}_f$ — среднее значение $\mathbf{f}$, а символом $|\mathbf{K}_f|$ обозначен определитель матрицы $\mathbf{K}_f$. Гауссова плотность является полезной моделью совместной плотности вероятности коэффициентов, полученных в результате унитарных преобразований изображений. Однако гауссова плотность не подходит для описания яркости элементов изображения, поскольку яркость может быть только положительной, а гауссовы случайные величины принимают как положительные, так и отрицательные значения.

В литературе редко встречаются выражения для совместных плотностей, не являющихся гауссовыми. Хунс [13] разработал методику формирования подобных плотностей высокого порядка на основе заданной безусловной плотности первого порядка и заданной ковариационной матрицы элементов ансамбля. Для плотности с нулевым средним эта процедура сводится к линейному преобразованию набора независимых случайных величин $\mathbf{g}$, совместную плотность вероятности которых

$$p(\mathbf{g}) = p\ [g(1)]\ p\ [g(2)] \cdots p\ [g(Q)] \qquad (5.5.4)$$

можно записать в виде произведения заданных плотностей первого порядка. Тогда искомая совместная плотность вероятности имеет вид

$$p_f(\mathbf{f}) = |\mathbf{A}|\, p_g(\mathbf{A}\mathbf{f}), \qquad (5.5.5)$$

где

$$\mathbf{A} = \mathbf{E}\boldsymbol{\Lambda}^{-1/2}\mathbf{E}^T, \qquad (5.5.6)$$

а $|\mathbf{A}|$ обозначает определитель матрицы $\mathbf{A}$. Столбцы матрицы $\mathbf{E}$ являются собственными векторами заданной ковариационной матрицы $\mathbf{K}_f$, а матрица $\boldsymbol{\Lambda}$ — диагональная и состоит из собственных значений этой матрицы, причем выполняется соотношение

$$\mathbf{E}^T \mathbf{K}_f \mathbf{E} = \boldsymbol{\Lambda}. \qquad (5.5.7)$$

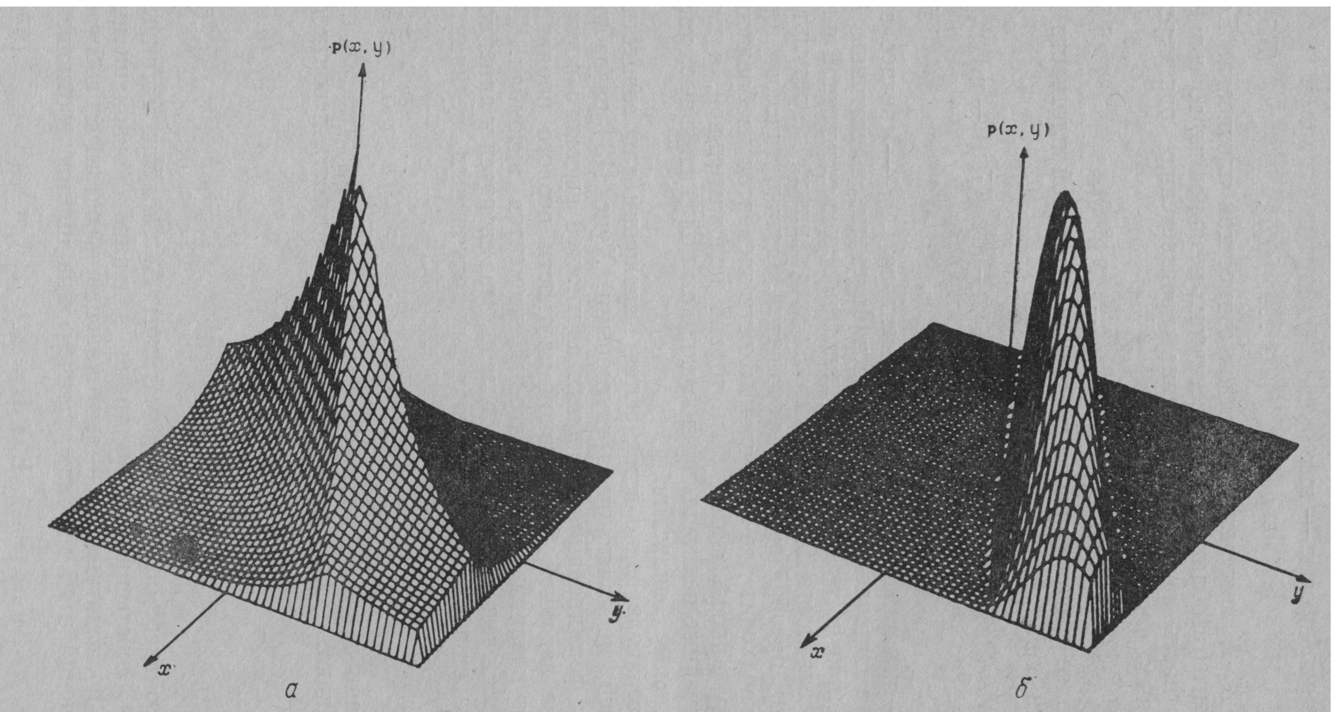


Рис. 5.5.1. Двумерные плотности вероятности пары коррелированных случайных величин ($\rho = 0{,}8$).
а — плотность распределения вероятностей Лапласа; *б* — плотность распределения вероятностей Рэлея.

На рис. 5.5.1 приведены двумерные плотности вероятности пары коррелированных случайных величин, безусловные распределения которых являются распределениями Рэлея или Лапласа. Многомерная модель с распределением Рэлея полезна при рассмотрении совместной плотности вероятности яркостей элементов изображения, а модель с распределением Лапласа применяется для статистического описания последовательности разностных сигналов, образующихся в системах кодирования изображений методом предсказания.

В следующей главе рассматривается методика квантования, т. е. представления отсчетов с помощью дискретного набора чисел, называемых уровнями квантования. Пусть $r_{j_q}(q)$ обозначает j_q-й уровень квантования для элемента изображения, который занимает q-е место в векторе изображения $\mathbf{f}$. Тогда вероятность получения одного из возможных значений вектора $\mathbf{f}$ можно выразить через совместное распределение вероятностей значений отсчетов следующим образом:

$$P(\mathbf{f}) = P_R\{f(1) = r_{j_1}(1),\ f(2) = r_{j_2}(2),\ \ldots,\ f(Q) = r_{j_Q}(Q)\}, \quad (5.5.8)$$

где $0 \leqslant j_q \leqslant J-1$. Обычно для всех компонент вектора выбирают одинаковый набор уровней квантования и совместное распределение вероятностей принимает вид

$$P(\mathbf{f}) = P_R\{f(1) = r_{j_1},\ f(2) = r_{j_2},\ \ldots,\ f(Q) = r_{j_Q}\}. \quad (5.5.9)$$

Распределения вероятностей значений отсчетов можно оценить, измеряя соответствующие частоты. Так, одномерное распределение q-й компоненты вектора

$$P\{f(q)\} = P_R\{f(q) = r_j\} \quad (5.5.10)$$

можно оценить путем анализа большого набора изображений, относящихся к одному и тому же классу, таких, например, как флюорограммы, аэрофотоснимки полей и т. д. Оценкой одномерного распределения вероятностей служит распределение относительных частот

$$H_E(j;\ q) = [N_p(j)]/N_p, \quad (5.5.11)$$

где N_p — полное число исследованных снимков, а $N_p(j)$ — число снимков, для которых $f(q) = r_j$, $j = 0, 1, \ldots, J-1$. Если источник изображений является стационарным, то одномерные распределения (5.5.10) будут одинаковы для всех компонент вектора, т. е. не будут зависеть от q. Кроме того, если источник изображений является эргодическим, то усреднение по ансамблю (измерения при использовании набора снимков) можно заменить усреднением по пространственным координатам. Если предполо-

жение об эргодичности справедливо, то одномерное распределение можно оценить по частотам

$$H_S(j) = N_S(j)/Q, \tag{5.5.12}$$

где $N_S(j)$ — число элементов исследуемого изображения, для которых $f(q) = r_j$, причем $1 \leqslant q \leqslant Q$, а $0 \leqslant j \leqslant J-1$.

На рис. 5.5.2 приведены одномерные гистограммы красной, зеленой и синей координат цвета цветного изображения «Портрет». В большинстве естественных изображений темных элементов

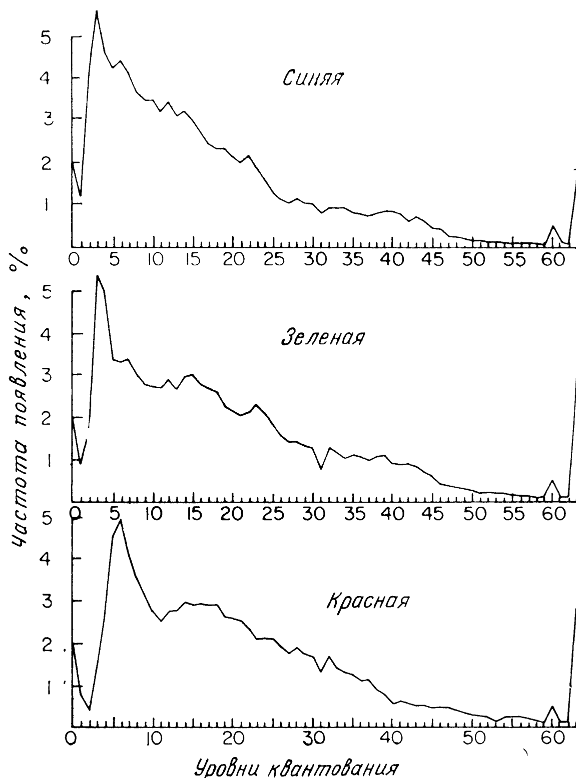


Рис. 5.5.2. Типичные гистограммы красной, зеленой и синей координат цвета цветного изображения.

гораздо больше, чем светлых, и частоты с ростом яркости убывают примерно по экспоненте.

Оценку двумерного распределения вероятностей для эргодического источника изображений можно найти с помощью распределения частот второго порядка, которые получают, подсчитывая случаи совместного появления определенных пар значений эле-

ментов, разделенных заданным расстоянием. Допустим, что $F(m_1, m_2)$ и $F(m_3, m_4)$ есть два элемента изображения, разделенные отрезком длиной r, который наклонен к горизонтальной

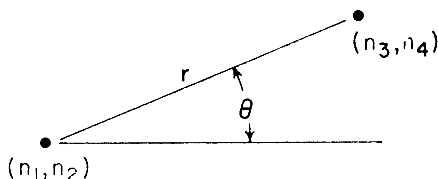


Рис 5.5 3. Взаимное расположение пары элементов изображения.

оси под углом θ (рис. 5.5.3). Поскольку элементы образуют прямоугольную решетку, параметры r, θ могут принимать только некоторые дискретные значения. Относительная частота второго порядка есть

$$H_S(j_1, j_2; r, \theta) = [N_S(j_1, j_2)]/Q_T, \tag{5.5.13}$$

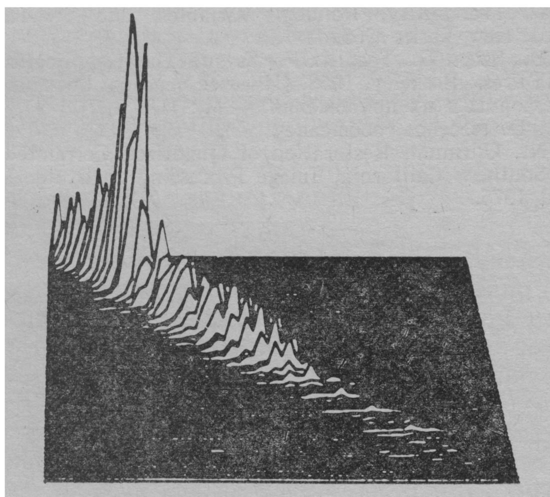


Рис. 5.5.4. Гистограммы второго порядка для изображения «Портрет».

где $N_S(j_1, j_2)$ — число пар элементов изображения, для которых $F(m_1, m_2) = r_{j_1}$ и $F(m_3, m_4) = r_{j_2}$. Знаменатель дроби (5.5.13) Q_T представляет собой полное число пар элементов, которые разделены отрезками с одинаковыми параметрами (r, θ). Из-за краевых эффектов $Q_T < Q$.

Гистограммы второго порядка для изображения «Портрет» представлены на рис.5.5.4. Элементы изображения с увеличением расстояния между ними становятся менее коррелированными, и частоты распределяются по плоскости более равномерно.

ЛИТЕРАТУРА

1. Ayres F., Jr., Schaum's Outline of Theory and Problems of Matrices, McGraw-Hill, New York, 1962.
2. Bellman R. E., Introduction to Matrix Analysis, McGraw-Hill, New York, 1970. [Имеется перевод: Беллман Р., Введение в теорию матриц. — М.: Наука, 1969.]
3. Campbell H. G., An Introduction to Matrices, Vectors, and Linear Programming, Appleton, New York, 1965.
4. Cullen C. G., Matrices and Linear Transformations, Addison-Wesley, Reading, Mass., 1966.
5. Graybill F. A., Introduction to Matrices with Applications in Statistics, Wadsworth, Belmont, Calif., 1969.
6. Rau C. R., Mitra S. K., Generalized Inverse of Matrices and its Applications, Wiley, New York, 1971.
7. Golub G. H., Reinsch C., Singular Value Decomposition and Least Squares Solutions, *Numer. Math.*, **14**, 403—420 (1970).
8. Andrews H. C., Patterson C. L., Outer Product Expansions and their Uses in Digital Image Processing, *Am. Math. Monthly*, **1**, 82, 1—13 (January 1975).
9. Pratt W. K., Vector Formulation of Two Dimensional Signal Processing Operations, *J. Comput. Graphics Image Proc.*, **4**, No. 1, 1—24 (March 1975) (Academic Press, New York).
10. Papoulis A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
11. Grenander U., Szego G., Toeplitz Forms and Their Applications, University of California Press, Berkeley, 1958. [Имеется перевод: Гренандер У., Сегё Г., Тёплицевы формы и их приложения. — М.: ИЛ, 1961.]
12. Davisson L. D. (частное сообщение).
13. Huhns M. N., Optimum Restoration of Quantized Correlated Signals, University of Southern California, Image Processing Institute, Report USCIPI 600, August 1975.

Глава 6

КВАНТОВАНИЕ ИЗОБРАЖЕНИЙ

Любая аналоговая величина, подлежащая обработке в ЦВМ или цифровой системе, должна быть представлена в виде целого числа, пропорционального значению этой величины. Процесс преобразования отсчетов, имеющих непрерывное множество значений, в отсчеты с дискретными значениями называется квантованием. В следующих двух разделах дан математический анализ процесса квантования, который справедлив не только для изображений, но и вообще для широкого класса сигналов, с которыми приходится сталкиваться в системах обработки изображений. Затем рассматривается обработка квантованных отсчетов. В двух последних разделах описаны субъективные эффекты, возникающие при квантовании одноцветных и цветных изображений.

6.1. КВАНТОВАНИЕ СКАЛЯРНЫХ ВЕЛИЧИН

Рис. 6.1.1 иллюстрирует типичный пример квантования скалярного сигнала. В процессе квантования значение отсчета аналогового сигнала сравнивается с набором пороговых уровней. Если отсчет попадает в интервал между двумя соседними пороговыми уровнями, то ему приписывается значение фиксированного уровня квантования, соответствующего данному интервалу. В цифровой системе каждому квантованному отсчету ставится в соответствие двоичная кодовая комбинация. В этом примере применен равномерный код, имеющий постоянную длину кодовых комбинаций.

Приступая к количественному анализу квантования скалярных величин, допустим, что f и $\hat{f}$ обозначают соответственно значения отсчета действительного скалярного сигнала до и после квантования. Предполагается, что f — случайная величина с плотностью вероятности $p(f)$. Кроме того, предполагается, что f не выходит за пределы некоторого интервала:

$$a_L \leqslant f \leqslant a_U, \qquad (6.1.1)$$

где a_U и a_L — верхняя и нижняя границы интервала.

При решении задачи о квантовании необходимо выбрать такой набор пороговых уровней d_j и уровней квантования r_j, что если

$$d_j \leqslant f \leqslant d_{j+1}, \tag{6.1.2}$$

то исходный отсчет заменяется на число, равное уровню квантования r_j. На рис. 6.1.2, *а* приведен пример размещения пороговых уровней и уровней квантования на отрезке числовой оси, содер-

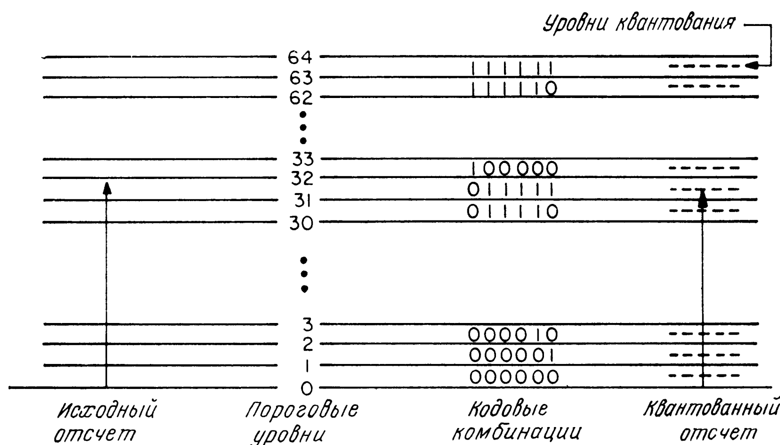


Рис. 6.1.1. Пример квантования сигнала.

жащем J пороговых уровней. Другой распространенной формой представления характеристики квантователя является ступенчатая кривая (рис. 6.1.2, *б*).

Уровни квантования и пороговые уровни выбирают так, чтобы уменьшить до минимума некоторую заданную величину, характеризующую ошибку квантования, т. е. степень различия между f и $\hat{f}$. В качестве меры ошибки квантования обычно выбирают среднеквадратическую ошибку. Если J — число уровней квантования, то среднеквадратическая ошибка квантования равна

$$\mathscr{E} = E\{(f - \hat{f})^2\} = \int_{a_L}^{a_U} (f - \hat{f})^2 p(f)\, df = \sum_{j=0}^{J-1} \int_{d_j}^{d_{j+1}} (f - r_j)^2 p(f)\, df. \tag{6.1.3}$$

Если число J велико, то плотность вероятности значений квантуемого сигнала на каждом из интервалов $(d_j,\ d_{j+1})$ можно считать постоянной и равной $p(r_j)$. Следовательно,

$$\mathscr{E} = \sum_{j=0}^{J-1} p(r_j) \int_{d_j}^{d_{j+1}} (f - r_j)^2\, df \tag{6.1.4}$$

или после вычисления интегралов

$$\mathcal{E} = \frac{1}{3} \sum_{j=0}^{J-1} p(r_j)\left[(d_{j+1} - r_j)^3 - (d_j - r_j)^3\right]. \tag{6.1.5}$$

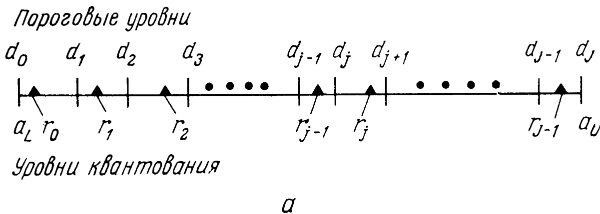
Оптимальное положение уровня квантования r_j в интервале $(d_j,\ d_{j+1})$ можно найти, решая задачу о минимуме ошибки $\mathcal{E}$ как функции r_j. Приравнивая нулю производную

$$d\mathcal{E}/dr_j = 0, \tag{6.1.6}$$

получаем

$$r_j = (d_{j+1} + d_j)/2. \tag{6.1.7}$$

Таким образом, при сделанных допущениях оптимальным положением уровня квантования является середина интервала между



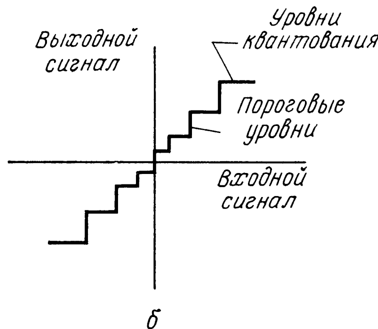


Рис. 6.1.2. Пороговые уровни и уровни квантования.

соседними пороговыми уровнями. Подставив соответствующие величины в выражение для ошибки квантования, получим

$$\mathcal{E} = 1/12 \sum_{j=0}^{J-1} p(r_j)\left[d_{j+1} - d_j\right]^3. \tag{6.1.8}$$

Оптимальное положение пороговых уровней можно определить, находя минимум ошибки $\mathcal{E}$ методом множителей Лагранжа. С помощью этого метода Пантер и Дайт [1] показали, что поло-

жения пороговых уровней довольно точно определяются по формуле

$$d_j = \frac{(a_U - a_L) \int\limits_{a_L}^{a_j} [p\{f\}]^{-1/3}\, df}{\int\limits_{a_L}^{a_U} [p\{f\}]^{-1/3}\, df} + a_L, \qquad (6.1.9а)$$

где

$$a_j = \frac{j\,(a_U - a_L)}{J} + a_L, \qquad (6.1.9б)$$

а $j = 0, 1, \ldots, J$. Если плотность вероятности значений отсчетов равномерна, то пороговые уровни будут расставлены равномерно. При неравномерных плотностях пороговые уровни чаще в тех участках, где плотность вероятности велика, и реже там, где она мала. Для большинства видов плотности вероятности, используемых обычно при описании изображений, интегралы (6.1.9) взять не удается и положение пороговых уровней приходится находить с помощью численного интегрирования.

Если число уровней квантования невелико, то приближение, с помощью которого получено равенство (6.1.4), становится неоправданным и следует использовать точное выражение для ошибки (6.1.3). Дифференцируя его по переменным d_j и r_j и приравнивая производные нулю, получаем

$$\partial\mathscr{E}/\partial d_j = (d_j - r_j)^2\, p\,(d_j) - (d_j - r_{j-1})^2\, p\,(d_j) = 0, \qquad (6.1.10а)$$

$$\partial\mathscr{E}/\partial r_j = 2 \int\limits_{d_j}^{d_{j+1}} (f - r_j)\, p\,(f)\, df = 0. \qquad (6.1.10б)$$

После преобразований приходим к системе уравнений

$$r_j = 2d_j - r_{j-1}, \qquad (6.1.11а)$$

$$r_j = \int\limits_{d_j}^{d_{j+1}} f_p\,(f)\, df \Bigg/ \int\limits_{d_j}^{d_{j+1}} p\,(f)\, df. \qquad (6.1.11б)$$

Решая эти уравнения рекуррентным способом, можно для заданной плотности вероятности $p\,(f)$ найти оптимальные значения пороговых уровней и уровней квантования. Макс [2] решил такую задачу для гауссовой плотности и составил таблицы оптимального размещения пороговых уровней в зависимости от числа уровней квантования. В табл. 6.1.1 указано расположение уровней квантования и пороговых уровней в квантователе Макса для плот-

Таблица 6.1.1

Расположение уровней квантования и пороговых уровней в квантователе Макса

Число разрядов	*Равномерное*		*Гаусса*		*Лапласа*		*Рэлея*	
	d_i	r_i	d_i	r_i	d_i	r_i	d_i	r_i
1	−1,0000	−0,5000	−∞	−0,7979	−∞	−0,7071	0,0000	1,2657
	0,0000	0,5000	0,0000	0,7979	0,0000	0,7071	2,0985	2,9313
	1,0000		∞		∞		∞	
2	−1,0000	−0,7500	−∞	−1,5104	−∞	−1,8340	0,0000	0,8079
	−0,5000	−0,2500	−0,9816	−0,4528	−1,1269	−0,4198	1,2545	1,7010
	−0,0000	0,2500	0,0000	0,4528	0,0000	0,4198	2,1667	2,6325
	0,5000	0,7500	0,9816	1,5104	1,1269	1,8340	3,2465	3,8604
	1,0000		∞		∞		∞	
3	−1,0000	−0,8750	−∞	−2,1519	−∞	−3,0867	0,0000	0,5016
	−0,7500	−0,6250	−1,7479	−1,3439	−2,3796	−1,6725	0,7619	1,0222
	−0,5000	−0,3750	−1,0500	−0,7560	−1,2527	−0,8330	1,2594	1,4966
	−0,2500	−0,1250	−0,5005	−0,2451	−0,5332	−0,2334	1,7327	1,9688
	0,0000	0,1250	0,0000	0,2451	0,0000	0,2334	2,2182	2,4675
	0,2500	0,3750	0,5005	0,7560	0,5332	0,8330	2,7476	3,0277
	0,5000	0,6250	1,0500	1,3439	1,2527	1,6725	3,3707	3,7137
	0,7500	0,8750	1,7479	2,1519	2,3796	3,0867	4,2124	4,7111
	1,0000		∞		∞		∞	
4	−1,0000	−0,9375	−∞	−2,7326	−∞	−4,4311	0,0000	0,3057
	−0,8750	−0,8125	−2,4008	−2,0690	−3,7240	−3,0169	0,4606	0,6156
	−0,7500	−0,6875	−1,8435	−1,6180	−2,5971	−2,1773	0,7509	0,8863
	−0,6250	−0,5625	−1,4371	−1,2562	−1,8776	−1,5778	1,0130	1,1397
	−0,5000	−0,4375	−1,0993	−0,9423	−1,3444	−1,1110	1,2624	1,3850
	−0,3750	−0,3125	−0,7995	−0,6568	−0,9198	−0,7287	1,5064	1,6277
	−0,2500	−0,1875	−0,5224	−0,3880	−0,5667	−0,4048	1,7499	1,8721
	−0,1250	−0,0625	−0,2582	−0,1284	−0,2664	−0,1240	1,9970	2,1220
	0,0000	0,0625	0,0000	0,1284	0,0000	0,1240	2,2517	2,3814
	0,1250	0,1875	0,2582	0,3880	0,2644	0,4048	2,5182	2,6550
	0,2500	0,3125	0,5224	0,6568	0,5667	0,7287	2,8021	2,9492
	0,3750	0,4375	0,7995	0,9423	0,9198	1,1110	3,1110	3,2729
	0,5000	0,5625	1,0993	1,2562	1,3444	1,5778	3,4566	3,6403
	0,6250	0,6875	1,4371	1,6180	1,8776	2,1773	3,8588	4,0772
	0,7500	0,8125	1,8435	2,0690	2,5971	3,0169	4,3579	4,6385
	0,8750	0,9375	2,4008	2,7326	3,7240	4,4311	5,0649	5,4913
	1,0000		∞		∞		∞	

ностей распределений вероятностей Гаусса, Лапласа, Рэлея и равномерного.

Нетрудно показать, что если пороговые уровни и уровни квантования выбраны согласно равенству (6.1.11), то среднеквадратическая ошибка квантования уменьшается до величины

$$\mathscr{E}_{\text{мин}} = \sum_{j=0}^{J-1} \left[\int_{d_j}^{d_{j+1}} f^2 p\{f\}\, df - r_j^2 \int_{d_j}^{d_{j+1}} p\{f\}\, df \right], \qquad (6.1.12а)$$

или в более короткой форме

$$\mathscr{E}_{\text{мин}} = E\{f^2\} - \sum_{j=0}^{J-1} r_j^2 P\{d_j \leqslant f < d_{j+1}\}. \qquad (6.1.12б)$$

Для частного случая плотности равномерного распределения вероятностей минимальная среднеквадратическая ошибка равна

$$\mathscr{E}_{\text{мин}} = 1/(12J^2). \qquad (6.1.13)$$

Для большинства других видов плотности вероятности ошибку квантования приходится определять с помощью расчетов.

Неравномерное квантование можно свести к равномерному с помощью нелинейного преобразования, как показано на

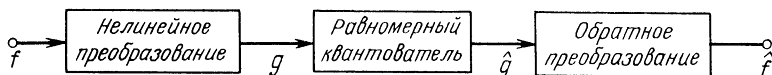


Рис. 6.1.3. Квантователь со сжатием.

рис. 6.1.3. Отсчет подвергается нелинейному преобразованию, затем равномерно квантуется и подвергается обратному нелинейному преобразованию [3]. В системах квантования с преобразованием стремятся сделать плотность вероятности преобразованных отсчетов на входе квантователя равномерной. Преобразованный отсчет (рис. 6.1.3) есть

$$g = T\{f\}, \qquad (6.1.14)$$

причем нелинейное преобразование $T(\cdot)$ выбрано таким, что плотность вероятности g оказывается равномерной, т. е.

$$p\{g\} = 1 \qquad (6.1.15)$$

в интервале $-1/2 \leqslant g \leqslant 1/2$. Если f — случайная величина с нулевым средним, то искомая характеристика нелинейного элемента имеет вид [4]

$$T\{f\} = \int_{\infty}^{f} p\{z\}\, dz - 1/2. \qquad (6.1.16)$$

Таблица 6.1.2

Квантование с преобразованием

Плотность вероятности	Прямое преобразование	Обратное преобразование
Гауссова $p(f) = (2\pi\sigma^2)^{-1/2} \times \exp - \{f^2/2\sigma^2\}$	$g = {}^1/_2 \operatorname{erf} \{f/\sqrt{2}\sigma\}$	$\hat{f} = \sqrt{2}\sigma \operatorname{erf}^{-1} \{2\hat{g}\}$
Рэлея $p(f) = (f/\sigma^2) \exp\{-f^2/2\sigma^2\}$	$g = {}^1/_2 - \exp\{-f^2/2\sigma^2\}$	$\hat{f} = [\sqrt{2}\sigma^2 \times \ln(1/({}^1/_2 - \hat{g}))]^{1/2}$
Лапласа $p(f) = (\alpha/2) \exp\{-\alpha\|f\|\}$, $\alpha = \sqrt{2}/\sigma$	$g = \begin{cases} {}^1/_2 (1 - \exp \cdot \{-\alpha f\}), & f \geqslant 0 \\ -{}^1/_2 (1 - \exp \cdot \{\alpha f\}), & f < 0 \end{cases}$	$\hat{f} = \begin{cases} -(1/\alpha) \times \ln(1 - 2\hat{g}), & \hat{g} \geqslant 0 \\ (1/\alpha) \times \ln(1 + 2\hat{g}), & \hat{g} < 0 \end{cases}$

$$\operatorname{erf}(x) = (2/\sqrt{\pi}) \int_0^x \exp(-y^2)\, dy$$

Таким образом, она совпадает с функцией распределения вероятностей величины f. В табл. 6.1.2 приведены характеристики прямого и обратного нелинейных преобразований для плотностей распределений вероятностей Гаусса, Релея, Лапласа.

6.2. КВАНТОВАНИЕ ВЕКТОРНЫХ ВЕЛИЧИН

Обычно квантование совокупности отсчетов выполняется последовательно. Каждый отсчет рассматривается как скалярная величина и квантуется независимо от остальных отсчетов с помощью методов, описанных в предыдущем разделе. Однако часто удается уменьшить ошибку квантования, если все отсчеты квантовать совместно.

Рассмотрим сигнал $\mathbf{f}$, представляющий собой вектор размера $N \times 1$. Будем полагать, что этот вектор является реализацией случайного вектора с N-мерной плотностью вероятности

$$p(\mathbf{f}) = p\{f_1, f_2, \ldots, f_N\}. \tag{6.2.1}$$

При квантовании вектора $\mathbf{f}$ N-мерное векторное пространство разделяют на J ячеек квантования D_j, каждой из которых соответствует один из J квантованных векторов. Векторный сигнал $\mathbf{f}$ заменяется на квантованный вектор $\mathbf{r}_j$, если $\mathbf{f}$ попадает в ячейку D_j. На рис. 6.2.1 приведены примеры векторного квантования в одно-, двух- и трехмерном пространствах. В подобной общей постановке задачи о векторном квантовании векторный сигнал $\mathbf{f}$ преобразуется в вектор $\mathbf{r}_j$, но компоненты вектора $\mathbf{f}$ при этом не обяза-

тельно будут квантоваться по отдельности по набору дискретных пороговых уровней.

Среднеквадратическую ошибку векторного квантования можно представить в виде суммы

$$\mathscr{E} = \sum_{j=0}^{J-1} \int_{D_j} \operatorname{tr}\left[(\mathbf{f}-\mathbf{r}_j)(\mathbf{f}-\mathbf{r}_j)^T\right] p(\mathbf{f})\, d\mathbf{f}. \qquad (6.2.2)$$

Оптимальное положение квантованных векторов $\mathbf{r}_j$ при фиксированных границах ячеек квантования D_j можно найти, приравнивая нулю частные производные ошибки квантования по векторам $\mathbf{r}_j$. В результате получается система интегральных уравнений

$$\partial\mathscr{E}/\partial\mathbf{r}_j = 0 = \int_{D_j} (\mathbf{f}-\mathbf{r}_j)\, p(\mathbf{f})\, d\mathbf{f}. \qquad (6.2.3)$$

После преобразований находим

$$\mathbf{r}_j = \int_{D_j} \mathbf{f} p(\mathbf{f})\, d\mathbf{f} \Big/ \int_{D_j} p(\mathbf{f})\, d\mathbf{f}. \qquad (6.2.4)$$

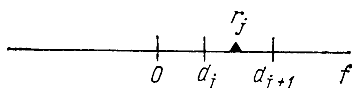
Равенство (6.2.4) определяет условное математическое ожидание вектора $\mathbf{f}$, когда он попадает в ячейку D_j:

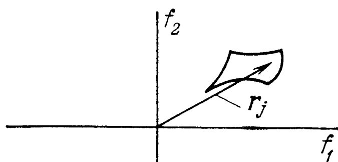
$$\mathbf{r}_j = E\{\mathbf{f} \mid \mathbf{f} \in D_j\}. \qquad (6.2.5)$$

В этом случае минимальная среднеквадратическая ошибка квантования равна

$$\mathscr{E}_{\text{мин}} = \operatorname{tr}\left[\mathbf{R}_f - \sum_{j=0}^{J-1} \mathbf{r}_j \mathbf{r}_j^T P\{\mathbf{f} \in D_j\}\right], \qquad (6.2.6)$$

где $\mathbf{R}_f$ — корреляционная матрица вектора $\mathbf{f}$. Заметим, что при $N = 1$ формула (6.2.4) сводится к (6.1.11), а выражение для ошибки квантования (6.2.6) переходит в формулу (6.1.12).





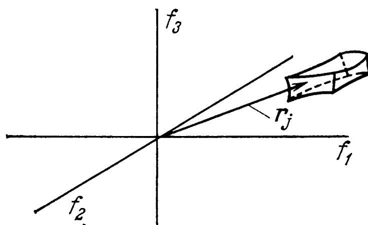


Рис. 6.2.1. Ячейки квантования в векторном пространстве.

а — одномерное пространство; *б* — двумерное пространство; *в* — трехмерное пространство

Оптимальные положения квантованных векторов $\mathbf{r}_j$ при фиксированных ячейках квантования D_j невозможно определить, не зная совместной плотности вероятности $p(\mathbf{f})$. Однако часто такая информация отсутствует. Еще одна существенная трудность связана с фактическим вычислением интегралов в формуле (6.2.4). Поэтому часто приходится упрощать процедуру векторного квантования. Так, можно квантовать все компоненты век-

тора $\mathbf{f}$ по отдельности, но задавать квантованные векторы $\mathbf{r}_j$ с помощью ячеек квантования D_j. Тогда в трехмерном пространстве ячейки D_j превращаются в прямоугольные параллелепипеды. Если при этом компоненты вектора $\mathbf{f}$ некоррелированы, то векторное квантование сводится к последовательному квантованию скалярных величин. Однако если отсчеты коррелированы, то задача определения оптимального вектора $\mathbf{r}_j$, как правило, не поддается решению без введения дополнительных упрощающих предположений. Кэри [4] получил решения применительно к совместным гауссовым плотностям, когда ячейки квантования D_j достаточно малы. Хунс [5] исследовал рекуррентный метод решения такой задачи, когда каждая компонента вектора определяется с помощью последовательных приближений на основе остальных квантованных компонент вектора. Этот метод позволяет найти решение задачи для множества различных плотностей вероятности: в части 6 он рассмотрен как средство уменьшения ошибок квантования в системах, использующих ИКМ и кодирование с преобразованием.

Попытаемся теперь найти такой набор ячеек квантования D_j, при котором минимизируется среднеквадратическая ошибка квантования. Брюс [6] разработал метод решения этой задачи, основанный на динамическом программировании. Однако в общем случае оптимальные формы ячеек квантования сложны и определить их весьма трудно. Поэтому в большинстве методов векторного квантования используется субоптимальный подход, когда для каждой компоненты вектора задается фиксированное число уровней квантования $J(i)$, где $i = 1, 2, \ldots, N$, и все компоненты квантуются независимо. Задача оптимизации сводится в этом случае к выбору величин $J(i)$, произведение которых

$$\prod_{i=1}^{N} J(i) = J \tag{6.2.7}$$

есть фиксированное число уровней квантования для данного вектора. Ошибка квантования i-го отсчета равна

$$\mathscr{E}(i) = E\{f^2(i)\} - \sum_{j=0}^{J(i)-1} r_j^2(i)\, P\{d_j(i) \leqslant f(i) < d_{j+1}(i)\}. \tag{6.2.8}$$

В системах с цифровым кодированием число уровней квантования обычно выбирают равным двоичному числу

$$J(i) = 2^{b(i)}, \tag{6.2.9}$$

где $b(i)$ — целое число кодовых разрядов (бит) для i-й компоненты вектора. Общее количество кодовых разрядов должно быть

постоянно и равно

$$B = \sum_{i=1}^{N} b(i). \tag{6.2.10}$$

Такой способ квантования называется блочным.

Несколько специалистов [7—9] разработали алгоритмы распределения числа разрядов $b(i)$ при фиксированном B, позволяющие минимизировать среднеквадратическую ошибку квантования. При использовании алгоритма, предложенного Реди и Уинцем [9] и применяемого для квантования независимых гауссовых величин по методу Макса, следует выполнить следующие операции:

1. Вычислить распределение числа разрядов по формуле

$$b(i) = (B/N) + 2\lg[\sigma^2(i)] - (2/N) \sum_{j=1}^{N} \lg[\sigma^2(j)], \tag{6.2.11}$$

где $\sigma^2(i)$ — дисперсия i-го отсчета.

2. Округлить каждое из чисел $b(i)$ до ближайшего целого.

3. Изменять полученное распределение, пока не будет выполнено условие (6.2.10)

Вывод соотношения (6.2.11) основан на экспоненциальной аппроксимации зависимости (6.1.12), связывающей ошибку квантования i-го отсчета и заданное ему число разрядов $b(i)$. Если $b(i)$ невелико, то эта аппроксимация оказывается довольно грубой. Более надежные результаты можно получить, пользуясь разработанным Прэттом [10] методом минимальной ошибки. Здесь разряды последовательно отводятся отсчетам, которые имеют наибольшую дифференциальную ошибку (6.2.8). При квантовании величин с нулевым средним алгоритм состоит из следующих шагов:

Шаг 1. Определение начальных условий:

B — общее число разрядов в кодовой комбинации блока,
N — длина блока,
$\sigma^2(n)$ — дисперсия компоненты,
$p\{f(n)\}$ — плотность вероятности компоненты,
B_s — индекс разряда (сначала равен нулю).

Шаг 2. Вычислить и запомнить коэффициенты дифференциальной ошибки

$$\tilde{D}[b(n)] \equiv \tilde{F}[2J(n)] - \tilde{F}[J(n)],$$

где $J(n) = 2^{b(n)}$, а сумма

$$\tilde{F}[J(n)] = \sum_{j=0}^{J(n)-1} \tilde{r}_j^2[J(n)] P\{\tilde{d}_j[J(n)] \leqslant$$

$$\leqslant f(n)/\sigma(n) < \tilde{d}_{j+1}[J(n)]\}$$

определяет коэффициент ошибки для случайных величин с единичной дисперсией, причем знаком тильды ($\sim$) отмечены уровни квантования и пороговые уровни, относящиеся к таким величинам.

Шаг 3. Отвести один разряд той компоненте, для которой произведение $\sigma^2(n)\,\tilde{D}\,[b\,(n)]$ имеет наибольшую величину; увеличить на единицу $b\,(n)$ и B_s.

Шаг 4. Если $B_s = B$, закончить процедуру; в противном случае повторить шаг 3.

Коэффициенты дифференциальной ошибки $\tilde{D}\,(\cdot)$, необходимые для данного алгоритма, можно вычислить заранее и хранить в виде таблицы из N чисел. В этом случае алгоритм сводится к B последовательным операциям распределения разрядов. Достоинства алгоритма состоят в том, что здесь не требуется громоздких вычислений, а также отсутствуют ошибки аппроксимации и округления, характерные для алгоритма с вычислением логарифмов дисперсий. Кроме того, метод минимальной ошибки опирается на модель плотности вероятности и поэтому его можно эффективно применять для распределения разрядов при неодинаковых плотностях вероятностей отсчетов.

6.3. ОБРАБОТКА КВАНТОВАННЫХ ВЕЛИЧИН

Числа, описывающие изображения (например, представляющие яркость или координаты цвета), обычно вводятся в цифровую вычислительную машину в виде целочисленных кодовых комбинаций, соответствующих уровням квантования отсчетов. Так, яркость одноцветного изображения обычно измеряют с помощью линейной целочисленной шкалы, ограниченной числами 0 (уровень черного) и 255 (уровень белого). Эти целочисленные кодовые комбинации, однако, не следует рассматривать как арифметические величины. Перед обработкой в машине кодовые комбинации следует преобразовать в действительные десятичные числа, соответствующие уровня квантования. Если этого не сделать, то можно получить совершенно неверные результаты. Так, если кодовые комбинации — номера уровней — изменяются вдоль шкалы яркостей немонотонно, то их вообще нельзя использовать для обработки. Рассмотрим теперь, что произойдет, если эти комбинации изменяются монотонно и обрабатываются в ЦВМ без преобразования в десятичные числа — значения уровней квантования.

Существуют две основные формы представления чисел в цифровых вычислительных машинах: как целое число и как действительное число. Целые числа изменяются от 0 до некоторого максимального значения. Так, в 16-разрядной мини-ЭВМ наибольшее положительное целое число равно 32 768 (2^{15}). Если результат арифметической операции определен как целое число, а в резуль-

тате получается дробное число, то дробная часть просто отбрасывается. Отношение 8/3, например, будет представлено как целое число 2 без десятичной точки и знаков после нее. Если обрабатываемые величины определены как действительные числа, то дробная часть итога операции будет сохраняться и содержать

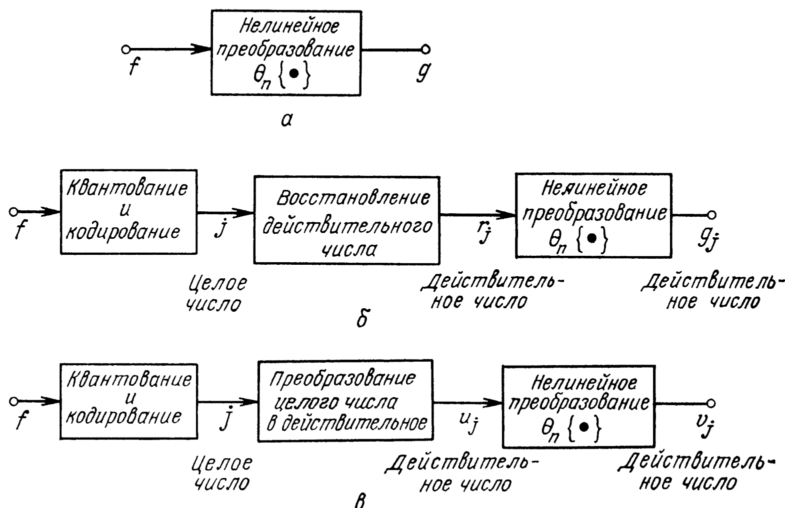


Рис. 6.3.1. Методы преобразования квантованных сигналов.

а — поэлементное преобразование непрерывных сигналов; *б* — преобразование квантованных сигналов, представленных действительными числами; *в* — преобразование квантованных сигналов при замене кодовых комбинаций действительными числами.

столько знаков, сколько позволяет получить разрядность машины. Отношение действительных чисел 8./3. будет иметь вид 2.66...66.

На рис. 6.3.1 для сравнения представлены три способа обработки сигналов. На рис. 6.3.1, *а* непрерывная скалярная величина f, изменяющаяся в интервале $a_L \leqslant f \leqslant a_U$, подвергается преобразованию $\mathcal{O}_p\{\cdot\}$, в результате которого получается непрерывная величина

$$g = \mathcal{O}_p\{f\}. \tag{6.3.1}$$

На рис. 6.3.1, *б* показана схема обработки, в которой скалярная величина f перед обработкой подвергается равномерному квантованию и кодированию. Целочисленные значения кодовых комбинаций определяются по формуле

$$j = \left[(J-1)\left(\frac{f-a_L}{a_U-a_L}\right)\right]_N, \tag{6.3.2}$$

где символ $[\cdot]_N$ обозначает округление аргумента до ближайшего целого числа. Код j, определенный как целое число, пре-

образуется в действительное число r_j согласно соотношению

$$r_j = [(r_{J-1} - r_0)/(J - 1)]\, j + r_0, \qquad (6.3.3а)$$

или, что то же самое,

$$r_j = [(a_U - a_L)/J]\, j + [(a_U - a_L)/2J] + a_L. \qquad (6.3.3б)$$

Далее с квантованным отсчетом r_j выполняется необходимая операция и получается квантованный выходной сигнал

$$g_j = \mathcal{O}_p\{r_j\}, \qquad (6.3.4)$$

отличающийся от непрерывного выходного сигнала g, фигурирующего в равенстве (6.3.1), только из-за ошибки квантования исходной величины.

К сожалению, часто обработку квантованных сигналов производят неправильно — по схеме, приведенной на рис. 6.3.1, *в*. Кодовая комбинация j, рассматриваемая как целое число, преобразуется в действительное число u_j, принимающее значения 0., 1., 2., ..., $(J - 1)$. . После этого по формуле

$$v_j = \mathcal{O}_p\{u_j\} \qquad (6.3.5)$$

вычисляются значения выходного сигнала v_j, определяемые как действительные числа. В общем случае v_j имеет и целую, и дробную части. Если, например $u_j = 17.$, то после преобразования, состоящего в извлечении квадратного корня с точностью до пятого знака, получается выходной сигнал $v_j = 4.12311$. Ясно, конечно, что ошибки квантования входного отсчета u_j проявятся и в выходном отсчете v_j. Однако существуют более серьезные трудности, если считать, что v_j может служить сравнительно хорошей аппроксимацией непрерывного выходного сигнала g. Предположим, что число уровней квантования достаточно велико и поэтому

$$u_j \approx (J - 1)\,[(f - a_L)/(a_U - a_L)]. \qquad (6.3.6)$$

Тогда выходной сигнал будет равен

$$v_j \approx \mathcal{O}_p\{k_1 f + k_2\}, \qquad (6.3.7)$$

где постоянные

$$k_1 = (J - 1)/(a_U - a_L), \qquad (6.3.8а)$$

$$k_2 = -[(J - 1)\, a_L]/(a_U - a_L). \qquad (6.3.8б)$$

Если преобразование $\mathcal{O}_p\{\cdot\}$ линейно, то

$$v_j \approx k_1 \mathcal{O}_p\{f\} + k_2 \qquad (6.3.9)$$

и

$$g \approx g_j \approx (v_j - k_2)/k_1. \qquad (6.3.10)$$

Таким образом, выходной сигнал рассматриваемой системы (рис. 6.3.1, *в*) будет хорошей аппроксимацией непрерывного

выходного сигнала системы с аналоговой обработкой (рис. 6.3.1, *а*), если преобразование, выполняемое в системе, является линейным. С другой стороны, если это преобразование нелинейно, то аппроксимация одного сигнала другим обычно оказывается очень плохой. Так, например, обычно логарифм квантованной переменной r_j существенно отличается от логарифма действительного числа u_j, сформированного из кодовой комбинации, которая задает номер уровня квантования для r_j.

6.4. КВАНТОВАНИЕ ОДНОЦВЕТНЫХ ИЗОБРАЖЕНИЙ

В системах передачи одноцветных изображений с применением импульсно-кодовой модуляции (ИКМ) каждый отсчет квантуется (обычно на основе линейной шкалы), и ему ставится в соответствие двоичная кодовая комбинация. Как правило, применяют равномерный код, и поэтому число уровней квантования светлоты выбирают из условия

$$L = 2^b, \tag{6.4.1}$$

где b — число двоичных разрядов (бит), отведенных для кодирования отсчетов.

В системах с ИКМ можно сократить цифровой поток простым уменьшением числа разрядов в кодовых комбинациях. Если имеется аналитическая мера качества изображений, передаваемых с помощью такой системы, то величина b будет равна наименьшему числу разрядов, при котором качество изображения еще является удовлетворительным. При субъективной оценке качества величину b понижают до тех пор, пока эффекты, вызванные квантованием, не выйдут за пределы допустимого. Глаз способен определить от 10 до 15 градаций абсолютного значения светлоты, однако он имеет гораздо большую чувствительность к различию светлоты соседних элементов изображения. При уменьшении числа уровней квантования прежде всего бросается в глаза эффект появления контуров в тех областях, где светлота исходного изображения изменяется плавно. Появление контуров вызвано скачкообразным изменением светлоты квантованного изображения при переходе от одного уровня квантования к другому. Наименьшее число разрядов ИКМ, позволяющее предотвратить появление контуров в областях с плавным изменением светлоты, зависит от ряда факторов, в том числе от линейности характеристик дисплея и свойств шума до и после преобразования изображения в видеосигнал.

Предположим, что отсчеты на выходе датчика видеосигнала пропорциональны яркостям элементов изображения. Возникают вопросы: следует ли квантовать непосредственно сигнал яркости изображения или же некоторую функцию от него и какой должна

быть характеристика квантователя — линейной или нелинейной? Последний вопрос скорее относится к практической реализации системы. Любую нелинейную характеристику квантователя можно получить, осуществляя нелинейное преобразование квантуемого сигнала, равномерное квантование и обратное нелинейное преобразование квантованного сигнала, как показано на схеме рис. 6.1.3. В связи с этим здесь будет рассматриваться только равномерное квантование отсчетов, которые предварительно могут быть подвергнуты нелинейному преобразованию.

Известно большое число экспериментальных работ по определению количества (и расположения) уровней квантования, необходимого для сведения к минимуму эффекта ложных контуров [11—14]. Гудолл [11] одним из первых проводил опыты с цифровым телевидением и пришел к выводу, что для получения хорошего качества необходимо квантовать яркостное изображение на 64 уровня (6 разрядов), а при 32 уровнях (5 разрядах) ложные контуры не слишком заметны. Аналогичные результаты получили и другие исследователи. В большинстве работ в той или иной мере затрагиваются вопросы о линейности характеристик системы воспроизведения изображений и ее калибровке. Сигналы телевизионной камеры и видеомонитора, как правило, нелинейно связаны с интенсивностью света. Фотоматериалы, приприменяемые для регистрации изображений, также имеют существенно нелинейные характеристики. Отметим, наконец, что любые шумы, создаваемые камерой или монитором, скрадывают ложные контуры.

На рис. 6.4.1 и 6.4.2 представлены фотографии, квантованные с разным числом уровней. На рис. 6.4.1 приведены изображения, полученные равномерным квантованием яркости при числе уровней от 2 до 64 (1÷ 6 двоичных разрядов). В темных частях изображения заметны ложные контуры, когда число разрядов не превышает пяти. На рис. 6.4.2 показаны результаты машинного моделирования процесса равномерного квантования оптической плотности изображения. В этом эксперименте величины $F(j, k)$ — квантованные на 256 уровней (8 разрядов) отсчеты яркости, значения которых лежат в диапазоне от 0 до 1, — были подвергнуты логарифмическому преобразованию:

$$G(j, k) = \{\ln [1 + F(j, k)\} / \ln 2. \qquad (6.4.2)$$

Величина $G(j, k)$ была равномерно квантована на разное число уровней. Функции $\hat{G}(j, k)$, описывающие изображение с квантованной плотностью, были затем преобразованы в функции

$$\hat{F}(j, k) = \exp\{\hat{G}(j, k) \ln 2\} - 1, \qquad (6.4.3)$$

которые описывают изображения с квантованной яркостью. Сравнение рис. 6.4.1 и 6.4.2 показывает, что при равномерном

Рис. 6.4.1. Пример равномерного квантования яркости изображения.

Рис. 6.4.2. Пример равномерного квантования оптической плотности изображения.

плотностном квантовании (рис. 6.4.2) эффект появления ложных контуров выражен не столь заметно, как при равномерном яркостном квантовании (рис. 6.4.1).

В гл. 19 рассматриваются методы цифрового кодирования изображений, предназначенные для подавления ложных контуров, появляющихся при недостаточном числе уровней квантования.

6.5. КВАНТОВАНИЕ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ

Цветное изображение можно описать с помощью красной, зеленой и синей координат цвета или произвольной (линейной или нелинейной) обратимой функции от этих величин. Если красная, зеленая и синяя координаты цвета квантуются по отдельности, то число и расположение уровней квантования выбираются по тем же общим правилам, что и при квантовании одноцветных изображений. Глаз реагирует на интенсивность света нелинейно, и поэтому субъективно качество квантованного цветного изображения оценивается выше, если перед квантованием подвергнуть координаты цвета нелинейному преобразованию, сжимающему их динамический диапазон. Известно, что глаз наиболее чувстви-

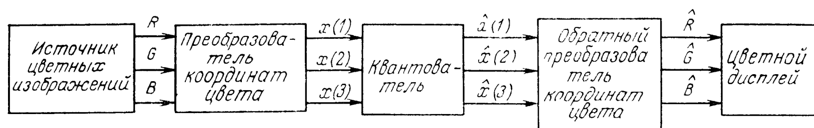


Рис. 6.5.1. Квантование цветных изображений.

телен к изменениям яркости в синей области спектра, умеренно чувствителен к таким изменениям в зеленой области и наименее чувствителен в красной области. Учитывая это, можно получить более эффективное распределение уровней квантования, чем равномерное.

На рис. 6.5.1 представлена обобщенная блок-схема системы квантования цветных изображений. Три координаты цвета R_N, G_N, B_N, описывающие исходное изображение, преобразуются в три компоненты вектора $x(1)$, $x(2)$, $x(3)$, которые и поступают в квантователь. Затем квантованные компоненты $\hat{x}(1)$, $\hat{x}(2)$ и $\hat{x}(3)$ преобразуются обратно в исходную координатную систему; в результате получается три координаты $\hat{R}_N$, $\hat{G}_N$, $\hat{B}_N$. Квантователь (рис. 6.5.2) по существу разделяет на ячейки квантования цветовое пространство, имеющее координаты $x(1)$, $x(2)$ и $x(3)$, и ставит в соответствие всем цветовым сигналам, попадающим в ячейку, одно и то же значение. Для большей эффективности следовало бы квантовать три цветовые компоненты $x(1)$, $x(2)$

и x (3) совместно. Однако по техническим причинам часто приходится квантовать цветовые компоненты независимо. В этом случае цветовое тело заключается в прямоугольный параллелепипед, который делится на прямоугольные ячейки квантования.

Если координаты цвета преобразуются в какую-то другую координатную систему, то сразу же возникают трудности. Рассмотрим, например, квантование координат цвета в системе UVW. На рис. 6.5.2 показано тело воспроизводимых цветов для системы

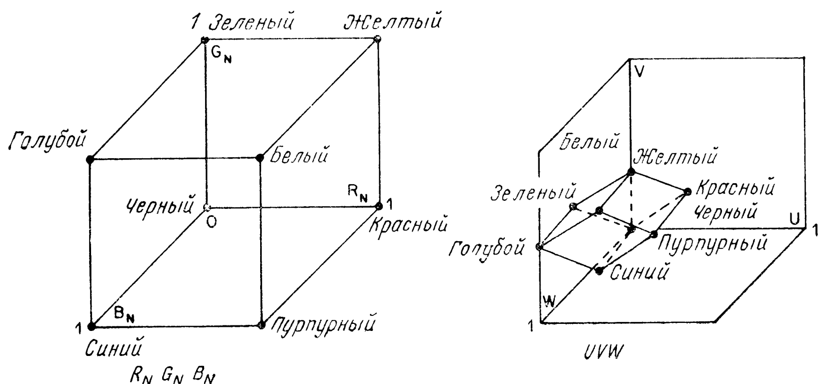


Рис. 6.5.2. Тело воспроизводимых цветов в системах координат $R_NG_NB_N$ и UVW.

$R_NG_NB_N$, представленное в форме куба, а также дан вид этого куба после преобразования его в координатную систему UVW. Можно заметить, что куб $R_NG_NB_N$ превращается в параллелепипед. Если координаты U, V, W квантуются независимо друг от друга, а диапазон квантования каждой из них охватывает весь интервал от минимального до максимального значения, то много ячеек квантования попадает в области невоспроизводимых цветов и поэтому оказываются бесполезными. Было бы желательно размещать все ячейки квантования внутри параллелепипеда воспроизводимых цветов, но, как правило, такую операцию сложно осуществить на практике.

Предположим для упрощения анализа, что каждая компонента квантуется равномерно на $2^{b(i)}$ уровней, причем $b(i)$ обозначает число разрядов, отведенных для квантования компоненты $x(i)$. Общее число таких разрядов фиксировано и равно

$$B = b(1) + b(2) + b(3). \tag{6.5.1}$$

Допустим, что $a_U(i)$ обозначает верхнюю, а $a_L(i)$ — нижнюю границу диапазона значений компоненты $x(i)$. Тогда размеры ячеек квантования равны

$$q(i) = [a_U(i) - a_L(i)]/2^{b(i)}. \tag{6.5.2}$$

Все цвета, цветовая компонента которых $x(i)$ попадает в одну и ту же ячейку, в ходе квантования будут заменяться одним цветом с квантованным значением $\hat{x}(i)$ этой компоненты. Ошибка квантования по каждой компоненте равна

$$\varepsilon(i) = |x(i) - \hat{x}(i)| = [a_U(i) - a_L(i)]/2^{b(i)+1}. \quad (6.5.3)$$

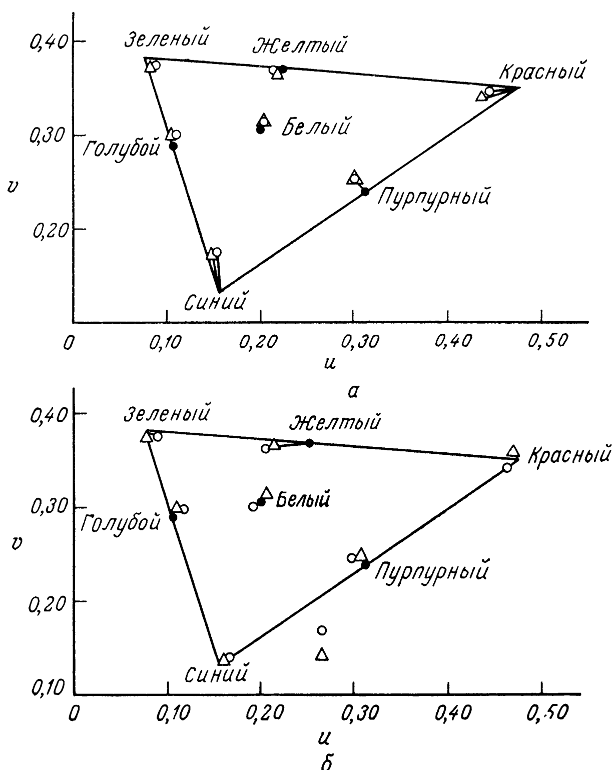


Рис. 6.5.3. Смещение цветов при равномерном квантовании изображения «Портрет» [15].

а — смещение цветов при квантовании в системе координат $R_N G_N B_N$; *б* — смещение цветов при квантовании в системе координат *Yuv*.
○ 4 разряда на каждую из трех координат;
△ 5, 4, 3 разряда на соответствующие координаты.

Тогда координата квантованного цвета

$$\hat{x}(i) = x(i) \pm \varepsilon(i), \quad (6.5.4)$$

причем $a_L(i) \leqslant \hat{x}(i) \leqslant a_U(i)$. Следует заметить, что значения $\hat{x}(i)$ всегда будут находиться внутри наименьшего куба, окружающего цветовое тело, для данной цветовой координатной системы.

На рис. 6.5.3 показаны смещения различных цветов при квантовании в системах координат $R_N G_N B_N$ и Yuv [15].

Джейн и Прэтт [15] рассмотрели вопрос об оптимальном распределении пороговых уровней при квантовании цветных изображений, позволяющем уменьшить до возможного предела цветовое расстояние между исходным цветом и его квантованным значением. Было установлено, что квантование в координатной системе $R_N G_N B_N$ дает лучшие результаты, чем в других распространенных цветовых координатных системах. Это объясняется тем, что в системе $R_N G_N B_N$ эффективно используются все ячейки квантования, а в других системах многие ячейки оказываются незанятыми. Последнее, по-видимому, существеннее, чем метрическая неоднородность цветового пространства $R_N G_N B_N$.

ЛИТЕРАТУРА

1. Panter P. F., Dite W., Quantization Distortion in Pulse Code Modulation with Non-uniform Spacing of Levels, *Proc. IRE*, **39**, 1, 44—48 (January 1951).
2. Max J., Quantizing for Minimum Distortion, *IRE Trans. Inf. Theory*, IT-6, 1, 7—12 (March 1960).
3. Algazi V. R., Useful Approximations to Optimum Quantization, *IEEE Trans. Commun. Tech.*, COM-**14**, 3, 297—301 (June 1966).
4. Curry R., Estimation and Control with Quantized Measurements, MIT Press, Cambridge, Mass., 1970.
5. Huhns M., Optimum Quantization Restoration, University of Southern California, Image Processing Institute Report 600, September 1975.
6. Bruce J. D., Optimum Quantization, MIT Research Laboratory of Electronics, Technical Report 429, March 1965.
7. Huang J. J. Y., Schutheiss P. M., Block Quantization of Correlated Gaussian Random Variables, *IEEE Trans. Commun. Syst.*, CS-**11**, 3, 289—296 (September 1963).
8. Ready P. J., Wintz P. A., Multispectral Data Compression Through Transform Coding and Block Quantization, Purdue University, Laboratory for Applications of Remote Sensing, Information Note 050572, May 1972.
9. Wintz P. A., Kurtenbach A. J., Waveform Error Control in PCM Telemetry, *IEEE Trans. Inf. Theory*, IT-**14**, 5, 650—661 (September 1968).
10. Pratt W. K., Block Quantization Bit Assignment (частное сообщение).
11. Goodall W. M., Television by Pulse Code Modulation, *Bell Syst. Tech. J.* (January 1951).
12. Cabrey R. L., Video Transmission over Telephone Cable Pairs by Pulse Code Modulation, *Proc. IRE*, **48**, 9, 1546—1551 (September 1960).
13. Harper L. H., PCM Picture Transmission, *IEEE Spectrum*, **3**, 6, 146 (June 1966).
14. Scoville F. .W, Huang T. S., The Subjective Effect of Spatial and Brightness Quantization in PCM Picture Transmission, NEREM Record, 1965, pp. 234—235.
15. Jain A. K., Pratt W. K., Color Image Quantization, National Telecommunications Conference 1972 Record, IEEE Publication No. 72 CHO 601-5-NTC, Houston, Texas, December 1972.

Глава 7

МЕРЫ КАЧЕСТВА ДИСКРЕТИЗОВАННЫХ ИЗОБРАЖЕНИЙ

При кодировании изображений для хранения или передачи с помощью минимально возможного количества цифр требуется сохранить качество воспроизводимого изображения в допустимых пределах. Для повышения визуального качества изображения разрабатываются системы улучшения изображений. Системы исправления изображений предназначаются для компенсации искажений и создания изображений, приближающихся к тем, которые были бы получены с помощью идеальной системы, не вносящей искажений. Во всех трех задачах общим является вопрос о качестве изображения — о том, чтобы в процессе обработки изображений поддерживать, повышать или восстанавливать его.

Качество изображения можно оценивать двояко: либо определяя верность воспроизведения, либо оценивая дешифрируемость изображения. Верность воспроизведения характеризует степень отклонения обработанного изображения от некоторого эталонного, а дешифрируемость указывает, в какой степени человек или машина может извлечь из изображения полезную информацию. Чаще всего верность воспроизведения связана с небольшими различиями между обработанным и эталонным изображениями. О дешифрируемости же, напротив, говорят, как правило, в связи с большими различиями между этими изображениями.

Очевидно, что количественные меры верности и дешифрируемости изображения очень нужны для проектирования и оценки систем воспроизведения изображений. Эти меры во многом помогут избавиться от трудоемкой и подчас неточной современной методики оценки изображений посредством субъективной экспертизы. Кроме того, на основе количественных мер можно развивать методы оптимизации систем обработки изображений.

В разработке количественных критериев верности и дешифрируемости изображений достигнуты значительные успехи. Однако введенные критерии не являются достаточно совершенными: очень часто можно привести примеры изображений, качество которых формально оценивается как высокое, а субъективно — как низкое и наоборот. Создание более совершенных критериев оценки качества изображений связано, несомненно, с более глубоким изучением свойств зрительной системы человека.

7.1. СУБЪЕКТИВНЫЕ ОЦЕНКИ КАЧЕСТВА ИЗОБРАЖЕНИЙ

В настоящее время наиболее распространенным и самым надежным способом определения качества изображений является субъективная экспертиза [1—7]. В некоторых случаях в качестве экспертов привлекают наблюдателей-неспециалистов. Их оценки определяют качество изображения именно так, как его воспринимает «средний» наблюдатель. Кроме того, проводятся опыты со специалистами, имеющими опыт обработки изображений, от которых следует ожидать более обоснованных оценок качества. Предполагается, что опытные наблюдатели замечают небольшие погрешности изображения, которые неспециалист может проглядеть.

Существуют два вида экспертных оценок: абсолютные и сравнительные. В первом случае наблюдатель должен оценить качество изображения по какой-то заранее определенной шкале. При некоторых методиках процесс оценивания облегчается тем, что наблюдателю предоставляется также набор эталонных изображений. Существуют и другие методики, когда наблюдатель вынужден принимать решение только на основании своего собственного опыта. При сравнительных оценках наблюдатель должен ранжировать набор конкретных изображений, т. е. расставить их в ряд по убыванию качества.

Распространенным видом шкалы оценок является шкала «общего качества», когда изображениям приписываются баллы, которые соответствуют градациям от «неудовлетворительно» до «отлично», приведенным в табл. 7.1.1 [1]. На практике каждое изображение поступает к наблюдателю, и он выставляет балл той категории, которая, по его мнению, в наибольшей степени соответствует качеству изображения. Вариантом данной шкалы является шкала «места в группе», или относительная шкала (табл. 7.1.2), когда наблюдатель оценивает качество изображения, сравнивая его со всеми другими изображениями, входящими в рассматриваемую группу [2]. Еще одной распространенной шкалой оценок является шкала погрешностей, согласно которой наблюдатель должен оценить в баллах степень искажений, изменяющихся от «незаметных» до «крайне нежелательных». Эта шкала, составленная Мерцем, Фаулером и Кристофером [3], приведена в табл. 7.1.3. Несколько других шкал оценки качества описаны в работе Пирсона [1].

Результаты экспертных оценок обычно выражают с помощью среднего балла, определяемого как

$$\bar{C} = \sum_{k=1}^{K} n_k C_k \Big/ \sum_{k=1}^{K} n_k, \tag{7.1.1}$$

где n_k — число изображений, отнесенных к k-й категории, а C_k — соответствующий ей балл. Считается, что для получения надеж-

ной оценки качества изображения необходимо опросить не менее двадцати наблюдателей. Одной из трудностей, связанных с балльными оценками, является возможная нелинейность шкалы. На рис. 7.1.1 для сравнения помещены шкала абсолютного (общего) качества (табл. 7.1.1), шкала погрешностей (табл. 7.1.3) и еще

Таблица 7.1.1

Шкала общего качества (абсолютная шкала)

5. Отлично
4. Хорошо
3. Посредственно
2. Плохо
1. Неудовлетворительно

Таблица 7.1.2

Шкала места в группе (относительная шкала)

7. Лучшее в группе
6. Заметно лучше среднего для данной группы
5. Несколько лучше среднего для данной группы
4. Среднее по группе
3. Несколько хуже среднего для данной группы
2. Существенно хуже среднего для данной группы
1. Самое плохое в группе

Таблица 7.1.3

Шкала погрешностей

1. Незаметные
2. Еле заметные
3. Вполне заметные, но слабо ухудшающие изображение
4. Ухудшающие изображение, но допустимые
5. Несколько нежелательные
6. Определенно нежелательные
7. Крайне нежелательные

одна шкала погрешностей, состоящая из трех градаций. Сравнение шкал выполнено на основе субъективных оценок [1].

Следует подчеркнуть, что на результаты субъективных экспертных оценок влияют характер рассматриваемых изображений и окружающая обстановка (условия эксперимента). Если наблюдатель видел подобные изображения, то он склонен более строго оценивать погрешности, поскольку имеет сложившееся представление о структуре изображения. С другой стороны, в незнакомом изображении искажения могут оставаться незамеченными, пока наблюдателю не будет на них указано. Ясно, что условия эксперимента должны в максимально возможной степени соответствовать условиям наблюдения в реальной обстановке. Кроме того,

следует с осторожностью пользоваться экспертными оценками, если изменились условия наблюдения. Так, качество изображения на экране обычного телевизионного монитора может быть расце-

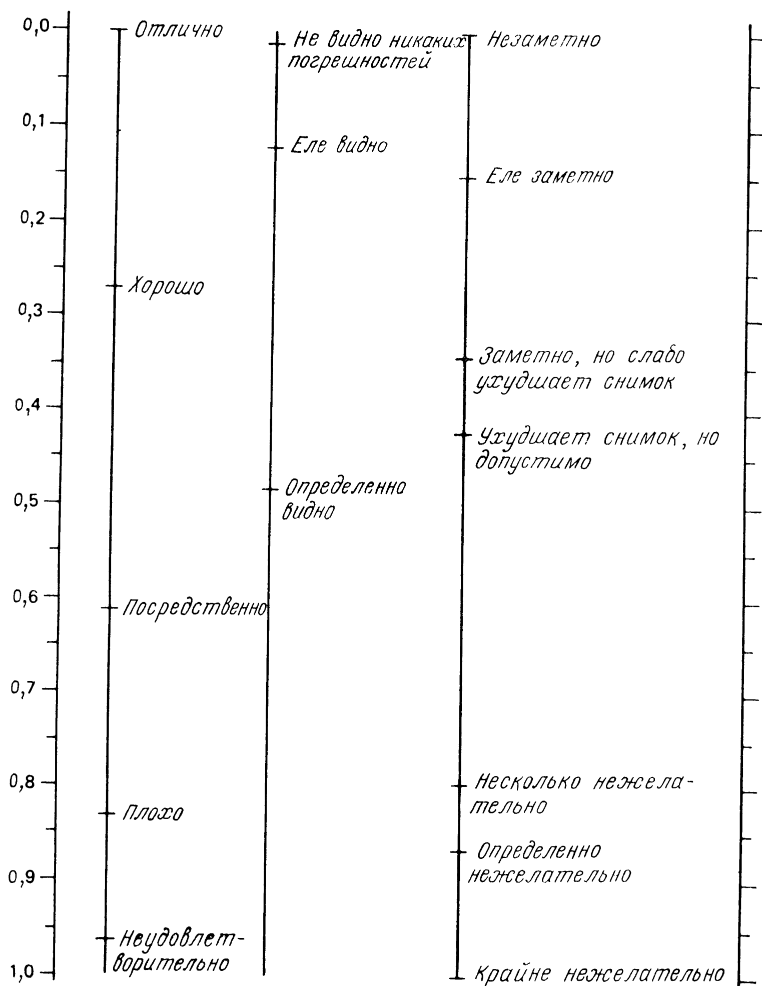


Рис. 7.1.1. Сравнение шкал качества и погрешностей, применяемых для субъективной оценки изображений.

нено как «хорошее», с «еле заметными» искажениями. Однако если то же самое изображение будет представлено в виде фотографии, снятой с помощью высококачественной аппаратуры, то все погрешности, ранее скрытые нелинейностями телевизионного устройства, неожиданно могут стать очень заметными.

7.2. ОЦЕНКА КАЧЕСТВА ПО КОНТРОЛЬНЫМ СИГНАЛАМ И ИСПЫТАТЕЛЬНЫМ ТАБЛИЦАМ

Телевизионные системы обычно проверяют и испытывают с помощью эталонных изображений, которые либо генерируются электронными устройствами, либо получаются обычным путем, когда перед камерой помещают испытательные таблицы. Этот

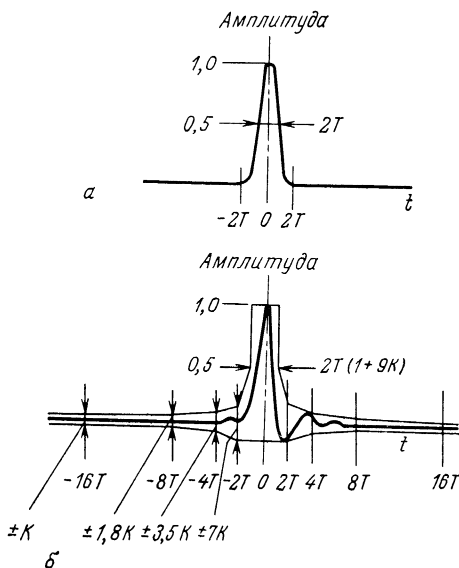


Рис. 7.2.1. Пример оценки искажений сигнала с помощью фактора *K* [8].

метод удобен благодаря его простоте, но связь результатов таких испытаний с результатами субъективной экспертизы, как правило, оказывается весьма сложной.

В аналоговых телевизионных системах искажение сигнала в канале связи часто оценивают сравнением переданного и принятого сигналов. Рис. 7.2.1 поясняет полуколичественную методику такого сравнения [8]. Испытательный сигнал (рис. 7.2.1, *а*) называют синус-квадратным импульсом. Принятый сигнал (рис. 7.2.1, *б*) заключают в рамку допусков, размеры которой определяются параметром *K*, и находят наименьшее значение параметра. Это значение выражает степень искажений.

С помощью испытательных электрических сигналов можно проверить наличие геометрических искажений, линейность передачи яркости и разрешающую способность телевизионного монитора. На рис. 7.2.2, *а* и 7.2.2, *б* помещены фотографии сформированных электронными средствами испытательных изображений,

применяемых для оценки геометрических искажений. Эти же изображения позволяют обнаружить несведение лучей в цветном телевизоре. Фотография шкалы градаций яркости, также сформированной электронными средствами, приведена на рис. 7.2.2, *в*. Линейность передачи яркостей можно оценить сравнением амплитуды сигнала с интенсивностью светового излучения от соответ-

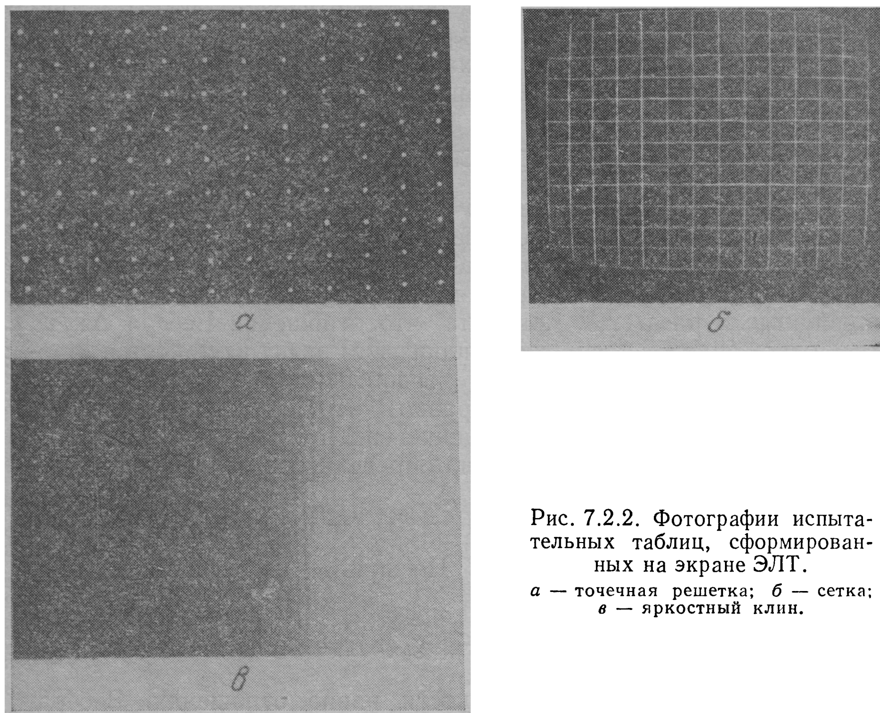


Рис. 7.2.2. Фотографии испытательных таблиц, сформированных на экране ЭЛТ.
а — точечная решетка; *б* — сетка; *в* — яркостный клин.

ствующих вертикальных полос на экране. Для проверки цветных телевизоров разработана таблица, состоящая из вертикальных полос белого, желтого, голубого, зеленого, пурпурного, красного и синего цветов [9—11].

Разрешающую способность можно определить, измеряя распределение интенсивности излучения в направлении, перпендикулярном границам полос.

Применение испытательных сигналов и таблиц позволяет измерить некоторые физические характеристики изображающих систем. Учитывая это, трудно переоценить полезность подобных методов для калибровки и настройки таких систем. Однако результаты испытаний по сигналам и таблицам не могут служить достаточно общими показателями качества изображения.

7.3. МЕТРИКИ СВЕТЛОТЫ И ЦВЕТА

Первый шаг в создании критерия верности воспроизведения изображения — это введение метрики для количественного измерения воспринимаемых различий в яркости, цветовом тоне и насыщенности точечного источника света. Существуют два вида мер различий характеристик света: глобальные и локальные. Глобальные меры пригодны для описания свойств любых источников света независимо от конкретных особенностей восприятия этого света. Локальные меры характеризуют только небольшие изменения параметров, описывающих световые ощущения.

7.3.1. МЕТРИКИ СВЕТЛОТЫ СВЕТА

Рассмотрим источник белого света $[L_1]$, имеющий яркость Y и расположенный в непосредственной близости от второго источника белого света $[L_2]$, имеющего такой же спектр излучения и яркость $Y + \Delta Y$. Измерения контрастной чувствительности, описанные в разд. 2.3, показали, что отношение Вебера $\Delta Y/Y$, т. е. отношение едва заметной (пороговой) разницы яркостей к абсолютному значению яркости Y, в довольно широком диапазоне значений Y почти постоянно и равно 1—2 %. На этом основании можно утверждать, что приращение светлоты второго источника по отношению к первому должно определяться логарифмическим выражением

$$\Delta\{L_1, L_2\} = \log_b(Y + \Delta Y) - \log_b(Y), \qquad (7.3.1)$$

а не разностью яркостей ΔY. Поскольку дифференциал логарифма яркости равен

$$d[\log_b(Y)] = \lim_{\Delta Y \to 0}[\log_b(Y + \Delta Y) - \log_b(Y)] = \log_b(e)(dY/Y), \qquad (7.3.2)$$

то приращение логарифма яркости равно отношению Вебера. Маннос и Сакрисон [2] изучали возможность измерения верности воспроизведения одноцветных изображений с помощью соотношения (2.4.8). В этом случае мерой приращения светлоты становится величина

$$\Delta\{L_1, L_2\} = (Y + \Delta Y)^{\nu} - (Y)^{\nu}, \qquad (7.3.3)$$

где ν — постоянная. В работе Корнсвита [12] было сделано предположение о том, что отклик фоторецепторов сетчатки глаза описывается нелинейным законом (2.4.7). Тогда соответствующая мера приращения светлоты

$$\Delta\{L_1, L_2\} = [K_1(Y + \Delta Y)]/[K_2 + (Y + \Delta Y)] - \\ - K_1 Y/(K_2 + Y), \qquad (7.3.4)$$

где K_1 и K_2 — постоянные.

Меры (7.3.1), (7.3.3) и (7.3.4) относятся к локальным мерам изменения светлоты. Часто используются и глобальные меры, которые служат для сравнения яркостей сильно различающихся источников света. Такие меры были получены на основе эмпирических моделей, построенных на экспериментальном материале. На рис. 7.3.1 приведены графики зависимости светлоты Λ от

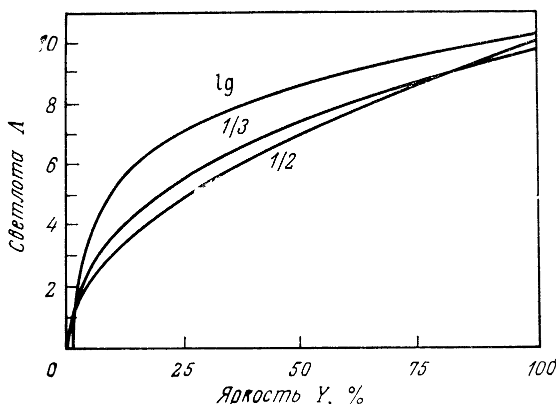


Рис. 7.3.1. Шкалы светлоты.

яркости источника белого света. Прист, Гибсон и Макниколс [13] предложили простое соотношение

$$\Lambda = (Y)^{1/2}, \tag{7.3.5}$$

где яркость измеряется в процентах ($0 \leqslant Y \leqslant 100$), а светлота изменяется от 0 до 10. Лэдд и Пинни [14] использовали шкалу с кубическим корнем

$$\Lambda = 2{,}468\,(Y)^{1/3} - 1{,}636, \tag{7.3.6}$$

где яркость также измеряется в процентах. В дополнение к этим зависимостям Фосс [15] ввел логарифмическую шкалу:

$$\Lambda = 5 \lg Y + 0{,}25. \tag{7.3 7}$$

Различие между тремя кривыми, приведенными на рис. 7.3.1, обусловлено главным образом различием уровней фонового освещения, применявшегося в соответствующих экспериментах [16, стр. 453]. Джадд [17] ввел шкалу светлоты, где учитывается уровень фонового освещения Y_B. В этом случае светлота определяется выражением

$$\Lambda = [0{,}1Y\,(Y_B + 100)]/(Y_B + Y). \tag{7.3.8}$$

Интересно отметить, что эмпирическая шкала Джадда по своей форме аналогична соотношению (2.4.7), описывающему нелинейную модель отклика фоторецепторов сетчатки.

7.3.2. МЕТРИКИ ЦВЕТА

Для определения цветовых различий можно ввести детерминированную меру, если рассматривать цветовое пространство, определяемое тремя координатами цвета, как евклидово пространство. Тогда цвет $[C_1]$, заданный значениями трех координат T_1, T_2 и T_3, будет отделен от цвета $[C_2]$ с координатами цвета $T_1 + \Delta T_1$, $T_2 + \Delta T_2$ и $T_3 + \Delta T_3$ расстоянием $\Delta\{C_1, C_2\}$, определяемым соотношением

$$\Delta^2\{C_1, C_2\} = (\Delta T_1)^2 + (\Delta T_2)^2 + (\Delta T_3)^2. \qquad (7.3.9)$$

Если данная мера различия цветов согласуется с субъективными представлениями, то при всех значениях координат цвета одинаковые приращения $\Delta\{C_1, C_2\}$ должны соответствовать одинаковым едва заметным (пороговым) различиям цветов. Другими словами, уравнение $\Delta^2\{C_1, C_2\} = 1$ должно описывать сферу радиуса в одну пороговую единицу различия цветов с центром в точке (T_1, T_2, T_3). Исключением являются только те точки, которые соответствуют чистым основным цветам — красному, зеленому и синему. Эксперименты показали, что человеческое зрение наиболее чувствительно к единичным изменениям синей координаты цвета, а наименее чувствительно — к зеленой [18].

Чтобы достигнуть лучшего соответствия меры (7.3.9) результатам субъективных испытаний, можно положить, что цветовое пространство является римановым. По определению расстояние в римановом цветовом пространстве выражается как

$$\Delta^2\{C_1, C_2\} = \sum_{i=1}^{3}\sum_{j=1}^{3} g_{ij}(\Delta T_i)(\Delta T_j), \qquad (7.3.10)$$

где коэффициенты g_{ij} являются функциями T_1, T_2, T_3. Приравняв выражение (7.3.10) единице, получим уравнение эллипсоида с центром в точке (T_1, T_2, T_3). Установлено, что цвета, едва заметно отличающиеся от данного цвета, можно довольно точно описать эллипсоидами в пространстве основных цветов.

Выражение (7.3.10) применяется в колориметрии в качестве меры различия цветов, если расстояние между ними не превышает нескольких пороговых единиц. Если же цвета различаются сильнее, то следует суммировать элементарные приращения вдоль кратчайшего пути, соединяющего соответствующие точки цветового пространства. Тогда глобальная мера различия цветов будет задана соотношением

$$d\{C_1, C_2\} = \underbrace{\sum_i \Delta_i\{C_1, C_2\}}_{\text{вдоль кратчайшего пути}}. \qquad (7.3.11)$$

В евклидовом пространстве таким кратчайшим путем является прямая линия. В римановом пространстве кратчайший путь, называемый геодезической линией, в общем случае проходит вдоль некоторой кривой. Мут и Перселс [19], а также Джейн [20] составили программы для вычисления расстояний между цветами, основанные на построении геодезических линий в пространстве цветов. На основе этих программ удалось создать весьма точную метрику цветового пространства, но для проведения вычислений требуется много времени.

Пользоваться римановой метрикой (7.3.10) сравнительно сложно, поэтому естественно возникает вопрос, существует ли такое отображение трехмерного пространства с координатами T_1, T_2, T_3 на евклидово трехмерное пространство с координатами $\tilde{T}_1$, $\tilde{T}_2$, T_3, при котором метрика различия цветов сохраняется неизменной. В этом новом пространстве приращения координат цвета должны удовлетворять соотношению

$$\Delta^2 \{C_1, C_2\} = (\Delta\tilde{T}_1)^2 + (\Delta\tilde{T}_2)^2 + (\Delta\tilde{T}_3)^2, \qquad (7.3.12)$$

так, чтобы поверхность едва различимых цветов представляла собой сферу с центром в точке $(\tilde{T}_1, \tilde{T}_2, \tilde{T}_3)$. Было показано [16, стр. 513], что в общем случае не существует линейного преобразования, в точности обладающего требуемыми свойствами. Однако было найдено несколько линейных преобразований со свойствами, близкими к заданным. Международная комиссия по освещению (МКО) приняла систему координат *UVW*, основанную на работах Джадда [21] и Мак-Адама [22]. В этой системе расстояния между цветами на пороге различения оказываются почти одинаковыми. На рис. 7.3.2, *а* приведены результаты измерений едва заметных отличий цветов, выполненных Мак-Адамом [23]. Множества едва отличимых цветов аппроксимированы эллипсами и представлены на графике цветностей в координатах x, y. Если эти же данные представить на графике цветностей с координатами u, v (рис. 7.3.2, *б*), то эллипсы становятся более близкими по размерам и заметно округляются. Фарнсуорт [24] нашел нелинейное преобразование, в результате которого эллипсы Мак-Адама (в пределах точности измерений Мак-Адама) превращаются в окружности (рис. 7.3.3). К сожалению, это преобразование выражается в виде сложной функции координат цветности x и y и требует выполнения громоздких вычислений.

В системе координат *UVW* расстояние между цветами равняется

$$\Delta^2 \{C_1, C_2\} = k_1 (\Delta U)^2 + k_2 (\Delta V)^2 + k_3 (\Delta)^2, \qquad (7.3.13)$$

где k_i — постоянные.

Человеческий глаз более чувствителен к изменению оттенков темных цветов. Стремясь использовать это обстоятельство, МКО

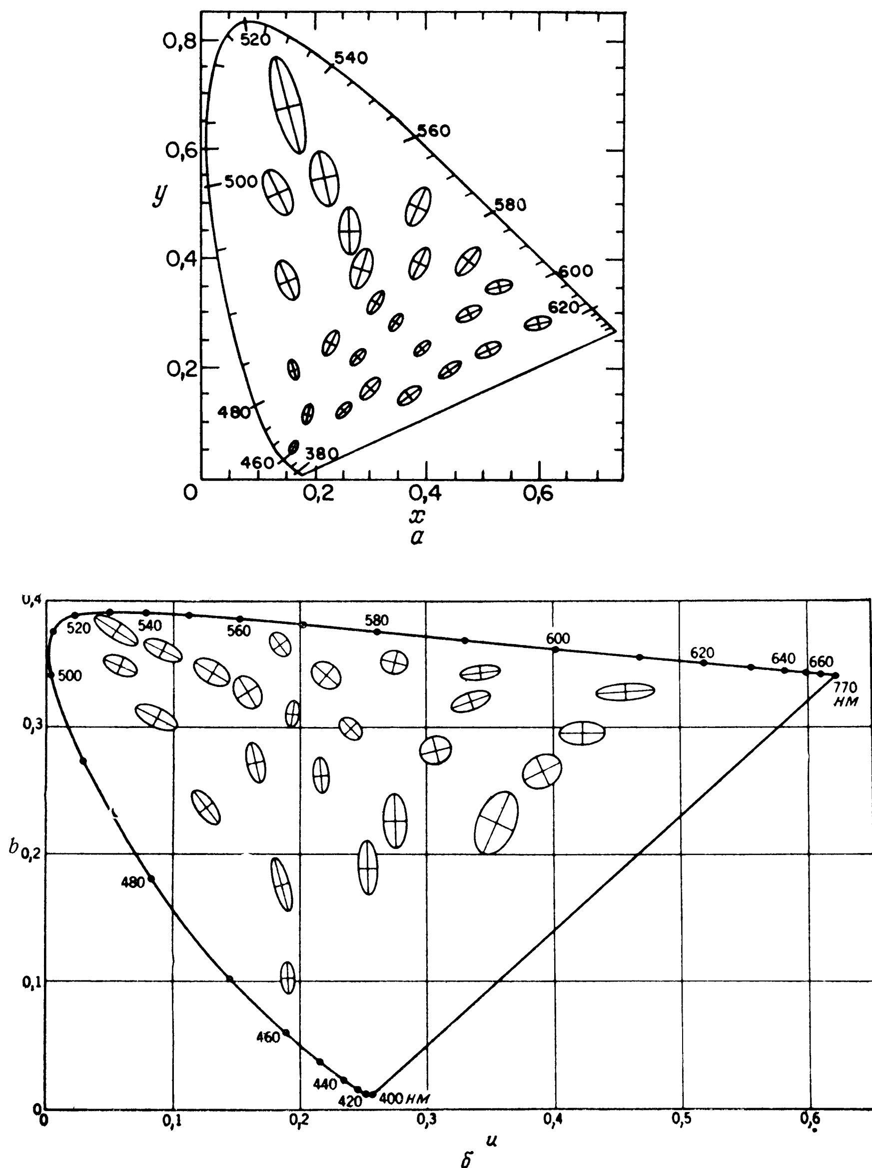


Рис. 7.3.2. Эллипсы Мак-Адама для пороговых различий цветов в системах координат *XYZ* и *UVW* [23]. Размеры эллипсов увеличены в 10 раз.

а — график цветностей в координатах *x*, *y*; *б* — график цветностей в координатах *u*, *v*.

временно ввела в качестве стандартов системы координат Lab и $U^*V^*W^*$, в которых изменения цветности и светлоты должны быть заметны в равной степени [25]. В этих системах различие цветов определяется величинами

$$\Delta^2 \{C_1, C_2\} = k_1 (\Delta L)^2 + k_2 (\Delta a)^2 + k_3 (\Delta b)^2, \quad (7.3.14а)$$

$$\Delta^2 \{C_1, C_2\} = k_1 (\Delta U^*)^2 + k_2 (\Delta V^*)^2 + k_3 (\Delta W^*)^2. \quad (7.3.14б)$$

Одним из недостатков меры различия цветов в системе $U^*V^*W^*$ является неразличимость разницы в цветовом тоне и насыщенности цвета. Считается, что различия в цветовом тоне гораздо более заметны, чем различия в насыщенности цвета. Поэтому была введена система координат $S\theta W$, в которой S пропорцио-

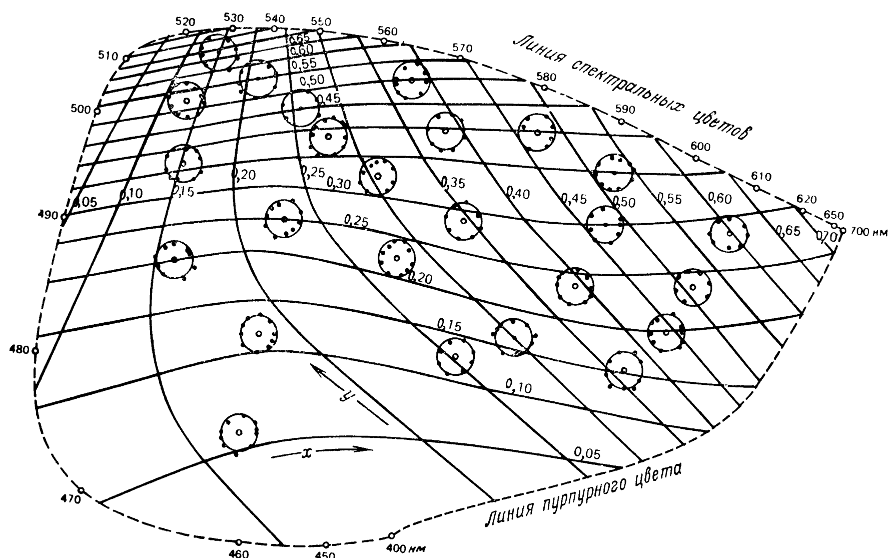


Рис. 7.3.3. Нелинейное преобразование эллипсов Мак-Адама на графике цветностей в системе координат XYZ [24].

нально насыщенности, а θ — цветовому тону [26, стр. 152]. Различие цветов в этой системе измеряется величиной

$$\Delta^2 \{C_1, C_2\} = k_1 (\Delta S)^2 + k_2 (\Delta\theta)^2 + k_3 (\Delta W^*)^2. \quad (7.3.15)$$

Если цвет имеет очень слабую насыщенность, т. е. близок к серому, то изменения цветового тона не имеют особого значения. На этом основании была введена еще одна мера различия цветов [26, стр. 152]:

$$\Delta^2 \{C_1, C_2\} = k_1 (\Delta S)^2 + k_2 [(S_1 + S_2)/2]^2 (\Delta\theta)^2 + k_3 (\Delta W^*)^2. \quad (7.3.16)$$

Все приведенные выше формулы являются результатами попыток создать субъективно-однородную меру цветовых различий. Экспериментов, которые позволили бы определить ценность каждой из них, проведено мало, так что вопрос о выборе подходящей меры все еще остается открытым.

7.4. ВЕРНОСТЬ ВОСПРОИЗВЕДЕНИЯ ОДНОЦВЕТНЫХ ИЗОБРАЖЕНИЙ

Построению и развитию системы количественных мер верности воспроизведения одноцветных изображений уделялось большое внимание [27—32]. Разумные меры верности должны хорошо согласовываться с результатами субъективных оценок для широкого класса изображений, не требуя при этом чересчур сложных вычислений. Кроме того, желательно, чтобы эти меры имели простую аналитическую форму и их можно было бы применять в качестве критериев оптимальности при оптимизации или выборе параметров систем обработки изображений.

Количественные меры верности воспроизведения одноцветных изображений можно разделить на две группы: одиночные и парные. Одиночная мера представляет собой число, сопоставляемое любому изображению на основе анализа его структуры. Парная мера является числовым результатом взаимного сравнения двух изображений.

Измерения верности в цифровой системе обработки изображений можно провести, используя либо непрерывное изображение, сформированное из массива отсчетов, либо сам этот массив. Обычно предпочитают второй способ, так как он проще с практической точки зрения. Однако, для того чтобы измерения на массиве отсчетов согласовывались с результатами субъективных оценок, воспроизводящее устройство не должно создавать больших или по крайней мере непредсказуемых искажений изображения. Ниже сначала будут введены меры верности изображений, получаемые на основе непрерывных двумерных функций, а затем представлены их дискретные варианты и описано соотношение между ними.

Рассмотрим непрерывную функцию $F(x, y)$, описывающую изображение, которая определена в прямоугольной области $-L_x \leqslant x \leqslant L_x$, $-L_y \leqslant y \leqslant L_y$. Предположим, что эта функция получена путем двумерной интерполяции массива отсчетов изображения $F(j, k)$ [при $-J \leqslant j \leqslant J$ и $-K \leqslant k \leqslant K$] согласно соотношению

$$F(x, y) = \sum_{j=-J}^{J} \sum_{k=-K}^{K} F(j, k) R(x - j\Delta x, y - k\Delta y), \quad (7.4.1)$$

где $R(x, y)$ — непрерывная интерполяционная функция, а Δx и Δy — шаги дискретизации.

Одиночную оценку верности в общем виде можно представить соотношением

$$Q = \int_{-L_x}^{L_x} \int_{-L_y}^{L_y} \mathcal{O}\{F(x, y)\}\, dx\, dy, \tag{7.4.2}$$

где $\mathcal{O}\{\cdot\}$ — некоторый (возможно, нелинейный) оператор. Критерии верности часто формулируют с использованием преобразований Фурье. В таком случае обобщенная форма одиночной меры верности имеет вид

$$Q = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{O}\{\mathcal{F}(\omega_x, \omega_y)\}\, d\omega_x\, d\omega_y, \tag{7.4.3}$$

где $\mathcal{F}(\omega_x, \omega_y)$ — непрерывный двумерный спектр Фурье изображения $F(x, y)$. Одна из наиболее простых мер подобного рода — предложенное Штрелем [33, стр. 461] отношение

$$Q = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}(\omega_x, \omega_y)\, d\omega_x\, d\omega_y}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}_I(\omega_x, \omega_y)\, d\omega_x\, d\omega_y}, \tag{7.4.4}$$

где $\mathcal{F}_I(\omega_x, \omega_y)$ — спектр исходного изображения. Первоначально это отношение использовалось для оценки качества элементов оптических систем. Спектральные составляющие оптического изображения в некоторой реальной оптической системе обычно уменьшаются по величине и, возможно, изменяются по фазе по сравнению с теми же составляющими в идеальной (безаберрационной) оптической системе, свойства которой ограничиваются только дифракционными явлениями. Составляющие с высокими пространственными частотами, как правило, ослабляются в наибольшей степени. С учетом определения преобразования Фурье (1.6.6а) отношение Штреля (7.4.4) сводится к $F(0, 0)/F_I(0, 0)$, т. е. к отношению центральных отсчетов реального и идеального изображений. Таким образом, отношение Штреля по существу является простой мерой уменьшения контраста реального изображения по сравнению с идеальным. Отношение Штреля в некоторой степени соответствует субъективным представлениям о качестве изображения, однако эксперименты показывают, что это соответствие не всегда полное. В частности, известны примеры изображений, которые обладали достаточно высокой дешифрируемостью, несмотря на то что отношение Штреля для них было небольшим [34].

Другим классическим примером одиночной меры верности изображения служит величина эквивалентного прямоугольника пропускания, равная по определению [35]

$$Q = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\mathscr{F}(\omega_x, \omega_y)|^2 \, d\omega_x \, d\omega_y. \qquad (7.4.5)$$

При возведении в квадрат возрастает вес составляющих изображения с низкими пространственными частотами, поскольку они, как правило, имеют большую величину. Однако и эта мера также не очень хорошо согласуется с результатами субъективных испытаний.

Попытки создания парных мер качества изображения имели несколько больший успех. Рассмотрим пару изображений, состоящую из некоторого эталонного (или идеального) изображения $F(x, y)$ и его искаженного варианта $\widehat{F}(x, y)$. Одной из мер «близости» двух изображений является их взаимная корреляция, равная по определению

$$\mathscr{K} = \int_{-L_x}^{L_x} \int_{-L_y}^{L_y} F(x, y) \widehat{F}(x, y) \, dx \, dy. \qquad (7.4.6)$$

Обычно взаимную корреляцию нормируют относительно энергии эталонного изображения так, чтобы ее максимальное значение равнялось единице. Нормированная взаимно-корреляционная мера — коэффициент корреляции — имеет вид

$$\mathscr{K} = \frac{\displaystyle\int_{-L_x}^{L_x} \int_{-L_y}^{L_y} F(x, y) \widehat{F}(x, y) \, dx \, dy}{\displaystyle\int_{-L}^{L_x} \int_{-L_y}^{L_y} [F(x, y)]^2 \, dx \, dy}. \qquad (7.4.7)$$

Согласно теореме Парсеваля (1.6.16), значения коэффициента корреляции можно вычислить по спектрам на основе соотношения

$$\mathscr{K} = \frac{\displaystyle\iint_{-\infty}^{\infty} \mathscr{F}(\omega_x, \omega_y) \widehat{\mathscr{F}}^*(\omega_x, \omega_y) \, d\omega_x \, d\omega_y}{\displaystyle\iint_{-\infty}^{\infty} |\mathscr{F}(\omega_x, \omega_y)|^2 \, d\omega_x \, d\omega_y}. \qquad (7.4.8)$$

При восприятии изображений важную роль играют контуры предметов. Поэтому Эндрюс [36] предложил пользоваться коэф-

фициентом корреляции лапласианов изображений, определяемым как

$$\mathscr{K} = \frac{\iint\limits_{-\infty}^{\infty} \left(\omega_x^2 + \omega_y^2\right) \mathscr{F}(\omega_x, \omega_y) \mathscr{F}^*(\omega_x, \omega_y)\, d\omega_x\, d\omega_y}{\iint\limits_{-\infty}^{\infty} \left(\omega_x^2 + \omega_y^2\right) |\mathscr{F}(\omega_x, \omega_y)|^2\, d\omega_x\, d\omega_y}. \qquad (7.4.9)$$

Напомним, что в силу соотношения (1.6.19) умножение спектра $\mathscr{F}(\omega_x, \omega_y)$ на квадрат частоты эквивалентно применению оператора Лапласа, что приводит к обострению контуров изображения, описываемого функцией $F(x, y)$. Эксперименты, выполненные Эндрюсом на изображениях, которые были преобразованы с помощью фильтров низких и высоких пространственных частот, показывают, что обычный коэффициент корреляции остается довольно большим даже в том случае, когда очень сильно подавлены высоко- и среднечастотные компоненты изображения и субъективно оно воспринимается как низкокачественное, а коэффициент корреляции лапласианов быстро уменьшается при сужении полосы низкочастотного фильтра. Можно, однако, получить низкокачественные изображения с большими искажениями в области низких пространственных частот, для которых коэффициент корреляции лапласианов оказывается сравнительно большим.

Еще одним парным критерием верности изображений является нормированная абсолютная ошибка — разность функций, описывающих эталонное и искаженное изображения:

$$\mathscr{E} = \frac{\int\limits_{-L_x}^{L_x} \int\limits_{-L_y}^{L_y} |\mathscr{O}\{F(x, y)\} - \mathscr{O}\{\hat{F}(x, y)\}|\, dx\, dy}{\int\limits_{-L_x}^{L_x} \int\limits_{-L_y}^{L_y} |\mathscr{O}\{F(x, y)\}|\, dx\, dy}. \qquad (7.4.10)$$

При обработке изображений в качестве меры ошибки чаще всего применяется нормированная среднеквадратическая ошибка, равная

$$\mathscr{E} = \frac{\int\limits_{-L_x}^{L_x} \int\limits_{-L_y}^{L_y} [\mathscr{O}\{F(x, y)\} - \mathscr{O}\{\hat{F}(x, y)\}]^2\, dx\, dy}{\int\limits_{-L_x}^{L_x} \int\limits_{-L_y}^{L_y} [\mathscr{O}\{F(x, y)\}]^2\, dx\, dy}. \qquad (7.4.11)$$

На практике обычно предпочитают пользоваться среднеквадратической, а не абсолютной ошибкой, поскольку первая удобнее для анализа, чем вторая. Именно поэтому проводились усиленные поиски таких преобразований, при которых среднеквадратическая ошибка преобразованной функции была бы хорошо согласована с субъективными оценками. В основном, конечно, рассматривались пространственные линейные и поэлементные нелинейные преобразования. Довольно много внимания уделялось степенным, а также логарифмическим преобразованиям. Исследовались операторы линейных пространственных преобразований, такие, как оператор градиента, оператор Лапласа и оператор свертки. Кроме того, рассматривались комбинации вышеупомянутых поэлементных и пространственных преобразований.

Выражение для ошибки можно также представить с помощью спектральных характеристик. В этом случае нормированная среднеквадратическая ошибка определяется соотношением

$$\mathscr{E} = \frac{\iint\limits_{-\infty}^{\infty} |\mathcal{O}\{\mathscr{F}(\omega_x, \omega_y)\} - \mathcal{O}\{\hat{\mathscr{F}}(\omega_x, \omega_y)\}|^2 \, d\omega_x \, d\omega_y}{\iint\limits_{-\infty}^{\infty} |\mathcal{O}\{\mathscr{F}(\omega_x, \omega_y)\}|^2 \, d\omega_x \, d\omega_y}. \tag{7.4.12}$$

Интересный частный случай линейного преобразования — это частотное взвешивание, когда

$$\mathcal{O}\{\mathscr{F}(\omega_x, \omega_y)\} = \mathscr{W}(\omega_x, \omega_y)\, \mathscr{F}(\omega_x, \omega_y), \tag{7.4.13}$$

где $\mathscr{W}(\omega_x, \omega_y)$ — весовая функция. Выражение для частотно-взвешенной среднеквадратической ошибки получается в виде

$$\mathscr{E} = \frac{\iint\limits_{-\infty}^{\infty} |\mathscr{W}(\omega_x, \omega_y)|^2 \, |\mathscr{F}(\omega_x, \omega_y) - \hat{\mathscr{F}}(\omega_x, \omega_y)|^2 \, d\omega_x \, d\omega_y}{\iint\limits_{-\infty}^{\infty} |\mathscr{W}(\omega_x, \omega_y)|^2 \, |\mathscr{F}(\omega_x, \omega_y)|^2 \, d\omega_x \, d\omega_y}. \tag{7.4.14}$$

Следует отметить, что формула (7.4.14) совершенно эквивалентна выражению (7.4.11) для среднеквадратической ошибки, если оператор $\mathcal{O}\{\cdot\}$ соответствует свертке, осуществляемой линейным фильтром с частотной характеристикой

$$\mathscr{H}(\omega_x, \omega_y) = |\mathscr{W}(\omega_x, \omega_y)|^2. \tag{7.4.15}$$

Уайлдер [37] провел глубокое исследование свойств абсолютной и среднеквадратической ошибок в дискретной форме применительно к степенному, логарифмическому и градиентному преобразованиям, а также к преобразованию Лапласа. Искаженные образцы изображений получались моделированием процесса

кодирования исходного изображения с помощью различных алгоритмов. Было выяснено, что поэлементные преобразования в сочетании с критериями абсолютной и среднеквадратической ошибок не позволяют получить критерия верности, согласующегося с субъективными оценками. Меры ошибки, основанные на преобразовании Лапласа и градиентном преобразовании, имеют наибольшую корреляцию с субъективными оценками, но коэффициент корреляции не превышает 0,8, т. е. и эти критерии недостаточно надежны.

Большинство попыток найти приемлемые критерии верности изображения относится к частным случаям. Предлагается некий критерий, основанный на каких-то физиологических предпосылках, а чаще просто удобный для анализа и вычислений, а затем оцениваются его свойства. Другой подход к проблеме состоит в копировании процесса выработки оценки человеком, т. е. свойства изображения должны измеряться в той метрике, которая присуща человеческому мозгу. При таком подходе оцениваемое изображение сначала проходит предварительную обработку и только потом оценивается его верность. При этом аппроксимируются, насколько это возможно, процессы, фактически происходящие в начальных звеньях зрительной системы человека. В гл. 2 была описана модель входного каскада зрительной системы человека, состоящая из трех звеньев. Двумерная линейная система с импульсным откликом $H_0(x, y)$ представляет оптические элементы глаза. Звено, выполняющее поэлементное нелинейное преобразование $\mathcal{O}_N\{\cdot\}$, моделирует отклик фоторецепторов. Вторая двумерная система с импульсным откликом $H_L(x, y)$ описывает процесс латерального торможения. Преобразование, выполняемое всей моделирующей системой, описывается выражением

$$\mathcal{O}\{F(x, y)\} = H_L(x, y) * \mathcal{O}_N\{H_0(x, y) * F(x, y)\}. \quad (7.4.16)$$

Маннос и Сакрисон [2] провели обширные измерения для разработки надежного среднеквадратического критерия верности одноцветных изображений, опирающегося на модель зрительной системы человека. Эффекты в оптической системе глаза, приводящие к ухудшению его разрешающей способности, при этом не учитывались, и соотношение (7.4.16) приводилось к более простому виду:

$$\mathcal{O}\{F(x, y)\} = H_L(x, y) * \mathcal{O}_N\{F(x, y)\}. \quad (7.4.17)$$

В этих экспериментах исходное распределение интенсивности, описывающее исходное изображение, подвергалось поэлементному нелинейному преобразованию по степенному или логарифмическому закону, а затем — пространственной фильтрации с частотной характеристикой

$$\mathcal{H}_L(\omega) = [c + (\omega/\omega_0)^{k_1}] \exp\{-(\omega/\omega_0)^{k_2}\}, \quad (7.4.18)$$

где c, k_1, k_2 и ω_0 — постоянные. Далее в полученное изображение вносились искажения, эквивалентные возникающим в процессе оптимального кодирования с заданным средним числом двоичных разрядов на элемент изображения. Затем искаженное изображение подвергалось обратной пространственной фильтрации с частотной характеристикой $\mathscr{H}_L^{-1}(\omega)$ и обратному поэлементному нелинейному преобразованию. В результате получалось искажен-

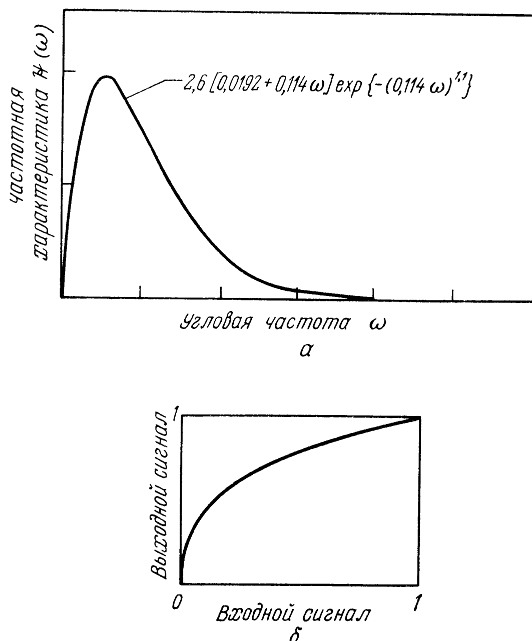


Рис. 7.4.1. Характеристики $\mathscr{H}_L(\omega)$ и $\mathscr{O}_N\{\cdot\}$, обеспечивающие наилучшее субъективное качество изображений при моделировании процесса кодирования [2].
а — частотная характеристика; *б* — характеристика поэлементного нелинейного преобразования.

ное изображение в виде распределения интенсивности. Все эти операции выполнялись над дискретным изображением. Подобная процедура повторялась для различных изображений при разных значениях среднего числа двоичных разрядов и других параметров. Качество полученных изображений оценивалось субъективно по семибалльной шкале места в группе (табл. 7.1.2) и ранжированием. Оказалось, что при одинаковом среднем числе двоичных разрядов на элемент наивысшие оценки и места в группе получили те изображения, которые подвергались действию фильтра с частотной характеристикой, приведенной на рис. 7.4.1. Кроме того, нелинейное преобразование по степенному закону с показателем $^1/_3$ дало гораздо лучшие результаты, чем логариф-

мическое преобразование. Исследования Манноса и Сакрисона показали также, что предварительная обработка согласно равенству (7.4.17), выполненная перед кодированием изображения, создает более благоприятные условия для кодирования. Результаты этих исследований удостоверяют еще и полезность применения среднеквадратического критерия верности с метрикой «геодезического пространства», характерного для человеческого зрения. Для количественного изучения этого критерия верности необходимы дальнейшие исследования.

В системах цифровой обработки изображений для определения верности воспроизведения обычно гораздо удобнее использовать дискретные отсчеты, а не аналоговые изображения. Поэтому важно найти критерии верности на основе дискретных отсчетов, которые хорошо согласуются с результатами субъективных оценок непрерывных изображений.

Непосредственный способ получения таких критериев верности состоит в простой «дискретизации» соответствующих аналоговых критериев. Так, например, нормированную среднеквадратическую ошибку (НСКО), описывающую различие между отсчетами $F(j, k)$ непрерывного эталонного изображения $F(x, y)$ и отсчетами $\hat{F}(j, k)$ непрерывного искаженного изображения $\hat{F}(x, y)$, можно представить как

$$\mathscr{E} = \sum_{j=1}^{J} \sum_{k=1}^{K} [\mathscr{O}\{F(j, k)\} - \mathscr{O}\{\hat{F}(j, k)\}]^2 \Bigg/ \sum_{j=1}^{J} \sum_{k=1}^{K} [\mathscr{O}\{F(j, k)\}]^2. \tag{7.4.19}$$

Можно показать, что эта мера совпадает с соответствующей «непрерывной» мерой, описываемой соотношением (7.4.11), если при дискретизации обоих изображений удовлетворяется критерий Найквиста. К сожалению, в реальных системах обработки изображений отсчеты, на основе которых создается воспроизводимое непрерывное изображение, не являются его найквистовскими отсчетами, так как воспроизводящее устройство вносит свои искажения и в нем трудно осуществить оптимальную двумерную интерполяцию. Тем не менее необходимость часто вынуждает пользоваться характеристиками, подобными (7.4.19), даже в тех случаях, когда они неточны.

В табл. 7.4.1 указаны некоторые наиболее распространенные критерии верности, основанные на оценках нормированных среднеквадратических ошибок (НСКО) для дискретных изображений. Еще одним широко применяемым критерием верности является так называемая пиковая среднеквадратическая ошибка (ПСКО), определяемая соотношением

$$\text{ПСКО} = (1/JK) \sum_{j=1}^{J} \sum_{k=1}^{K} [G(j, k) - \hat{G}(j, k)]^2 / A^2, \tag{7.4.20}$$

где $G(j, k)$ — преобразованное изображение, соответствующее определению, приведенному в табл. 7.4.1, а число A равно максимальному значению $G(j, k)$. Среднеквадратические ошибки часто

Таблица 7.4.1

Критерии верности, основанные на оценках нормированных среднеквадратических ошибок, для дискретных одноцветных изображений

Без преобразования

$$\text{НСКО} = \sum_{j=1}^{J} \sum_{k=1}^{K} [F(j, k) - \hat{F}(j, k)]^2 / \sum_{j=1}^{J} \sum_{k=1}^{K} [F(j, k)]^2$$

С поэлементным преобразованием

$$\text{НСКО} = \sum_{j=1}^{J} \sum_{k=1}^{K} [G(j, k) - \hat{G}(j, k)]^2 / \sum_{j=1}^{J} \sum_{k=1}^{K} [G(j, k)]^2$$

Степенной закон

$$G(j, k) = [F(j, k)]^{\nu}$$

Логарифмический закон

$$G(j, k) = k_1 \log_b [k_2 + k_3 F(j, k)]$$

С преобразованием Лапласа

$$\text{НСКО} = \sum_{j=2}^{J-1} \sum_{k=2}^{K-1} [G(j, k) - \hat{G}(j, k)]^2 / \sum_{j=2}^{J-1} \sum_{k=2}^{K-1} [G(j, k)]^2,$$

где

$$G(j, k) = F(j+1, k) + F(j-1, k) + F(j, k+1) + F(j, k-1) - 4F(j, k)$$

С использованием свертки

$$\text{НСКО} = \sum_{j=1}^{J} \sum_{k=1}^{K} [G(j, k) - \hat{G}(j, k)]^2 / \sum_{j=1}^{J} \sum_{k=1}^{K} [G(j, k)]^2,$$

где

$$G(j, k) = [F(x, y) * H(x, y)]\, \delta(x - j\,\Delta x,\ y - k\Delta y)$$

измеряют в децибелах и рассматривают как отношение сигнал/шум (С/Ш):

$$(\text{С/Ш})_{\text{Н}} = -10 \lg (\text{НСКО}) \tag{7.4.21а}$$

или

$$(\text{С/Ш})_{\text{П}} = -10 \lg (\text{ПСКО}). \tag{7.4.21б}$$

Имея дело с операторами пространственных преобразований, такими, как оператор Лапласа или оператор свертки (табл. 7.4.1),

не следует забывать, что массивы отсчетов изображений имеют конечные размеры. Поэтому пределы суммирования [в выражениях, подобных (7.4.19) или (7.4.20)] необходимо ограничивать центральными областями массивов $G(j, k)$ и $\hat{G}(j, k)$ с тем, чтобы избежать краевых эффектов, которые появляются при аппроксимации непрерывных интегралов свертки дискретными суммами. Более подробно эти вопросы рассматриваются в гл. 11.

7.5. ВЕРНОСТЬ ВОСПРОИЗВЕДЕНИЯ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ

Разработка количественных критериев верности воспроизведения цветных изображений является гораздо более трудной задачей, чем создание таких критериев для одноцветных изображений. Это обусловлено не только увеличением числа величин, описывающих изображение, но и необходимостью учета большего числа психологических факторов, связанных с восприятием цветов. Критерий верности воспроизведения цветных изображений должен согласовываться с мерой различия цветов и такими явлениями, как цветные полосы Маха и цветовая адаптация.

В разд. 7.3.2 было получено выражение для меры различия цветов

$$\Delta^2\{C_1, C_2\} = \sum_{i=1}^{3}\sum_{j=1}^{3} g_{ij}\,\Delta T_i\,\Delta T_j, \tag{7.5.1}$$

где ΔT_i обозначает разность координат цвета, а g_{ij} — весовые коэффициенты, зависящие от координат цвета. Исходя из этой формулы, примем за критерий верности воспроизведения цветных изображений нормированную среднеквадратическую ошибку. По определению эта мера равна

$$\mathcal{E} = \frac{\iint\limits_{-\infty}^{\infty} \sum\limits_{i=1}^{3}\sum\limits_{j=1}^{3} g_{ij}\,[T_i(x, y) - \hat{T}_i(x, y)]\,[T_j(x, y) - \hat{T}_j(x, y)]\,dx\,dy}{\iint\limits_{-\infty}^{\infty} \sum\limits_{i=1}^{3}\sum\limits_{j=1}^{3} g_{ij} T_i(x, y)\,T_j(x, y)\,dx\,dy}, \tag{7.5.2}$$

где $T_i(x, y)$ и $\hat{T}_i(x, y)$ — цветовые компоненты эталонного и искаженного изображений соответственно. Среднеквадратическую ошибку довольно трудно вычислять по этой формуле, так как в нее входят перекрестные члены — произведения различных координат цвета. Кроме того, весовые коэффициенты g_{ij} не постоянны, а зависят от координат цвета. Как было отмечено при

рассмотрении вопроса о метрике цвета, не существует такого линейного преобразования, при котором перекрестные члены исчезали бы полностью. С помощью нелинейных преобразований координат, описанных в гл. 3, осуществляется переход к цветовому пространству $U^*V^*W^*$ или Lab, временно принятым МКО в качестве стандартных пространств с равномерным распределением яркости и цветности. В пространстве $U^*V^*W^*$, например, выражение для среднеквадратической ошибки имеет вид

$$\mathscr{E} = \frac{\displaystyle\iint_{-\infty}^{\infty} \left(h_1 [U^*(x, y) - \hat{U}^*(x, y)]^2 + h_2 [V^*(x, y) - \hat{V}^*(x, y)]^2 + h_3 [W^*(x, y) - \hat{W}^*(x, y)]^2\right) dx\, dy}{\displaystyle\iint_{-\infty}^{\infty} \left(h_1 [U^*(x, y)]^2 + h_2 [V^*(x, y)]^2 + h_3 [W^*(x, y)]^2\right) dx\, dy}. \qquad (7.5.3)$$

При другом подходе к выбору среднеквадратической меры верности воспроизведения цветного изображения используются некоррелированные между собой координаты цвета, полученные в результате преобразования первичных координат. Такую декорреляцию можно провести с помощью разложения Карунена—Лоэва, описанного в гл. 3. Исчезновение перекрестных членов в полученной таким образом мере интуитивно обосновывается тем, что координаты цвета являются статистически некоррелированными. Преобразование Карунена—Лоэва определяется ковариационной матрицей координат цвета $T_i(x, y)$ исходного изображения, которую необходимо либо измерить (оценить), либо получить из модели данного класса цветных изображений. Система координат YIQ оказалась довольно близкой к системе Карунена—Лоэва [38]. По этой причине среднеквадратическую ошибку, вычисленную для координат Y, I, Q, часто применяют как меру верности воспроизведения цветных изображений. К сожалению, пока нет достаточных данных, чтобы судить о согласованности мер, рассмотренных в этом разделе, с результатами субъективных оценок.

Третий способ определения верности воспроизведения цветных изображений состоит в вычислении среднеквадратической ошибки в цветовом геодезическом пространстве, согласующемся с некоторой моделью цветового зрения человека. Для подобной модели, приведенной на рис. 2.5.3, среднеквадратическая ошибка имеет вид

$$\mathscr{E} = \iint_{-\infty}^{\infty} \sum_{i=1}^{3} [g_i(x, y) - \hat{g}_i(x, y)]^2\, dx\, dy \Bigg/ \iint_{-\infty}^{\infty} \sum_{i=1}^{3} [g_i(x, y)]^2\, dx\, dy, \qquad (7.5.4)$$

где $g_i(x, y)$ — сигналы, получаемые из сигналов колбочек $e_i(x, y)$ с помощью следующих преобразований:

$$g_1(x, y) = H_1(x, y) * \log[e_1(x, y)], \qquad (7.5.5а)$$

$$g_2(x, y) = H_2(x, y) * \log\left[\frac{e_2(x, y)}{e_1(x, y)}\right], \qquad (7.5.5б)$$

$$g_3(x, y) = H_3(x, y) * \log\left[\frac{e_3(x, y)}{e_1(x, y)}\right]. \qquad (7.5.5в)$$

Переменные $e_i(x, y)$ в свою очередь связаны с координатами цвета $T_i(x, y)$ некоторым линейным преобразованием

$$\mathbf{e} = \mathbf{MT}, \qquad (7.5.6)$$

где **M** — матрица преобразования размера 3×3. Как и в предыдущих случаях, экспериментальная проверка пригодности геодезической меры ошибки почти не проводилась.

Приведенные выше меры верности воспроизведения непрерывных цветных изображений можно приспособить и для оценки дискретных цветных изображений, пользуясь методикой, примененной для дискретизации мер верности воспроизведения одноцветных изображений.

7.6. ЭНТРОПИЯ ИЗОБРАЖЕНИЙ

Очевидно, что одни изображения являются более содержательными, чем другие, т. е. имеют больше деталей или при анализе из них удается извлечь больше данных. Детали, данные и другие подобные понятия являются качественными и довольно расплывчатыми. Поэтому часто бывает необходимо ввести количественные характеристики изображения, позволяющие оценивать предельные свойства алгоритмов кодирования, исправления и анализа изображений. Один из подходов к количественному описанию изображений состоит в применении теории информации [39—42].

Согласно методике, описанной в разд. 5.3, допустим, что матрица **F** размера $N \times N$, составленная из квантованных отсчетов изображения, будет заменена вектор-столбцом **f** размера $Q \times 1$ (где $Q = N^2$), полученным путем развертки матрицы **F** по столбцам. В принципе такие векторы можно рассматривать как изображения на выходе некоторого источника, который может генерировать любой из возможных векторов. В одном из крайних случаев получается темное изображение с минимальной яркостью всех элементов, а в противоположном случае — изображение с максимальной яркостью. Между этими крайними случаями заключено множество различных изображений. Если яркость каждого из $N \times N$ элементов изображения квантуется на J уровней, то данный источник может создать $T = J^Q$ различных изобра-

жений. Многие из них имеют хаотическую структуру и похожи на реализации двумерного случайного шума. Лишь очень небольшое число из T возможных изображений будут такими, какие дал бы реальный датчик при наблюдении за окружающим миром. В принципе можно считать, что имеется априорное распределение $P(\mathbf{f}_t)$ (где $t = 1, 2, \ldots, T$) вероятности появления каждого из возможных состояний вектора $\mathbf{f}$. Измерить или промоделировать это распределение очень непросто, но сама идея в конце концов приводит к полезным результатам.

В 1948 г. Шеннон [39] опубликовал свою знаменитую книгу «Математическая теория связи» в которой был дан способ количественного описания свойств источников данных и систем передачи информации. В основе шенноновской теории информации лежит понятие об энтропии. При векторном описании изображения среднее количество информации в изображении равно энтропии источника:

$$H(\mathbf{f}) = -\sum_{t=1}^{T} P(\mathbf{f}_t) \log_2 P(\mathbf{f}_t). \tag{7.6.1}$$

В данном определении энтропии применяются логарифмы с основанием два и энтропия измеряется в двоичных единицах. Энтропию источника полезно знать при кодировании изображений, поскольку согласно теореме о кодировании при отсутствии помех [39], теоретически можно закодировать без искажений изображения, создаваемые источником с энтропией $H(\mathbf{f})$, затратив на кодирование $H(\mathbf{f}) + \varepsilon$ двоичных единиц, где ε — бесконечно малая положительная величина. И наоборот, в принципе невозможно закодировать изображения без искажений, если число двоичных единиц меньше чем $H(\mathbf{f})$.

Вероятность появления t-го вектора, описывающего изображение, можно выразить через совместное распределение вероятностей уровней яркости элементов изображения

$$P(\mathbf{f}_t) = \Pr\{f(1) = r_{j1}(1),\ f(2) = r_{j2}(2),\ \ldots,\ f(Q) = r_{jQ}(Q)\}, \tag{7.6.2}$$

где $r_{jq}(q)$ — значение j-го уровня квантования для q-го элемента [1]. Эту вероятность можно выразить также в виде произведения условных вероятностей:

$$\begin{aligned} P(\mathbf{f}) = \Pr\{f(1) = r_{j1}(1)\} \Pr\{f(2) = r_{j2}(2) \mid f(1)\} \Pr\{f(3) = \\ = r_{j3}(3) \mid f(2),\ f(1)\} \cdots \Pr\{f(q) = \\ = r_{jq}(q) \mid f(q-1),\ f(q-2),\ \ldots,\ f(2),\ f(1)\} \cdots \Pr\{f(Q) = \\ = r_{jQ}(Q) \mid f(Q-1),\ f(Q-2),\ \ldots,\ f(2),\ f(1)\}. \end{aligned} \tag{7.6.3}$$

---

[1] Предполагается, что источник изображений не обладает памятью в том смысле, что распределение вероятностей для данного изображения не зависит от того, какие изображения были созданы источником ранее.

Логарифмируя обе части равенства (7.6.3) по основанию два и учитывая определение энтропии (7.6.1), получаем

$$H(\mathbf{f}) = -\sum_{t=1}^{T} P(\mathbf{f}_t)\log[\Pr\{f(1) = r_{j1}(1)\}] -$$
$$-\sum_{t=1}^{T} P(\mathbf{f}_t)\log[\Pr\{f(2) = r_{j2}(2)\,|\,f(1)\}] - \cdots -$$
$$-\sum_{t=1}^{T} P(\mathbf{f}_t)\log[\Pr\{f(q) = r_{jq}(q)\,|\,f(q-1),\ \ldots,\ f(1)\}] \ldots . \tag{7.6.4}$$

В равенстве (7.6.4) q-е слагаемое, обозначенное через $H[f(q)|$ $|f(q-1),\ \ldots,\ f(1)]$, можно рассматривать как выражение для средней информации, содержащейся в q-й компоненте вектор-изображения $\mathbf{f}$, при условии, что точно известны яркости предшествующих $q-1$ компонент. Таким образом,

$$H(\mathbf{f}) = \sum_{q=1}^{Q} H[f(q)\,|\,f(q-1),\ \ldots,\ f(1)]. \tag{7.6.5}$$

Это выражение, описывающее энтропию источника изображений, является общим и не зависит от того, в каком порядке берутся элементы изображения. Рассмотрим теперь вид формулы (7.6.5) для двух случаев: 1) изображение развертывается по столбцам и 2) когда все элементы изображения поступают одновременно. Можно показать, что при развертке изображения по столбцам

$$H[f(q)\,|\,f(q-1),\ \ldots\ f(q-j)] \geqslant H[f(q)\,|\,f(q-1),\ \ldots,\ f(q-k)], \tag{7.6.6}$$

где $j > k$. Это означает, что информация, содержащаяся в q-й точке, с увеличением числа полностью известных предшествующих элементов изображения в среднем только уменьшается. Равенство обеих частей соотношения (7.6.6) достигается только тогда, когда яркости всех элементов распределены независимо [1]. Если количество известных предшествующих элементов неограниченно увеличивается, то правая часть неравенства (7.6.6) стремится к некоторому ненулевому предельному значению, обозначаемому как $H[f(q)\,|\,\infty]$.

Если пренебречь малосущественными для достаточно крупных изображений краевыми эффектами, то энтропию изображения можно приближенно выразить в виде

$$H(\mathbf{f}) \approx QH[f(q)\,|\,\infty]. \tag{7.6.7}$$

[1] Равенство в соотношении (7.6.6) имеет место в том случае, если последовательность элементов, полученная в результате развертки, является марковской последовательностью не более чем k-го порядка. При этом нет необходимости неограниченно увеличивать число предшествующих элементов для получения величины $H[f(q)\,|\infty]$. — *Прим. ред.*

Таким образом, можно считать, что энтропия всего изображения равна предельному значению условной энтропии одного элемента изображения, умноженному на полное число элементов.

В системах с разверткой изображения предельная условная энтропия определяется на основе конечной последовательности предыдущих элементов. Так, если этих элементов было k, то

$$H\,[f(q)\,|\,\infty] \approx H\,[f(q)\,|\,f(q-1),\ \ldots,\ f(q-k)]. \qquad (7.6.8)$$

В явной форме

$$H\,[f(q)\,|\,f(q-1),\ \ldots,\ f(q-k)] = \sum_{j_0=0}^{J-1} \cdots \sum_{j_k=0}^{J-1} P\,[f(q),\ \ldots,\ f(q-k)] \log \left\{ \frac{P\,[f(q),\ \ldots,\ f(q-k)]}{P\,[f(q-1),\ \ldots,\ f(q-k)]} \right\}, \qquad (7.6.9)$$

где совместное распределение вероятностей

$$P\,[f(q),\ \ldots,\ f(q-k)] = \Pr\,\{f(q) = r_{j0}(q),\ \ldots,\ f(q-k) = r_{jk}(q-k)\} \qquad (7.6.10)$$

(r_{jq} — уровни квантования яркости). Для вычисления условной энтропии (7.6.9) необходимо либо ввести модель совместного распределения вероятностей, либо измерить соответствующие распределения частот для изображения или некоторого класса изображений.

Таблица 7.6.1

Оценки энтропии изображения, полученные Шрайбером

Порядок	Функциональное выражение энтропии	Энтропия, бит/элемент
Первый	$H\,[f(q)]$	4,4
Второй	$H\,[f(q)\,\|\,f(q-1)]$	1,9
Третий	$H\,[f(q)\,\|\,f(q-1),\ f(q-2)]$	1,5

Шрайбер [43] оценил энтропию первого, второго и третьего порядка для нескольких изображений, квантованных на 64 уровня, измеряя распределения относительных частот того же порядка. Полученные распределения были подставлены в формулу (7.6.10) вместо соответствующих распределений вероятностей. При таком способе измерений предполагается, что источник изображений является стационарным и эргодическим, т. е. усреднение по ансамблю изображений можно заменить усреднением по отдельному изображению. Результаты измерений, проведенных Шрайбером для конкретного изображения, приведены в табл. 7.6.1. Для кодирования этого изображения с помощью обычной ИКМ

требуются шестиразрядные кодовые слова, т. е. затрачивается 6 бит/элемент. Теоретически для его кодирования достаточно 4,4 бит/элемент при условии, что все элементы кодируются по отдельности. Если же использовать значение предыдущего элемента, то теоретический предел уменьшится до 1,9 бит/элемент. Иначе говоря, предыдущий элемент дает 2,5 бита информации об элементе $f(q)$. Учет еще одного предшествующего элемента дает всего 0,4 бита дополнительной информации. Таким образом, оказалось, что в данном изображении бóльшая часть дополнительно извлекаемой информации содержится в небольшом числе предыдущих элементов.

Выше был описан метод оценки энтропии изображения, развернутого по столбцам, при котором предельная условная энтропия $H[f(q)\,|\,\infty]$ аппроксимируется условной энтропией, вычисленной с учетом нескольких предыдущих элементов в том же столбце. Такой метод можно применить для оценки энтропии изображения, когда все его элементы поступают одновременно.

Допустим, что $q = (Q+1)/2$ обозначает номер центрального элемента вектора $\mathbf{f}$ с нечетным числом элементов. Тогда совместное распределение вероятностей можно представить в виде

$$P(\mathbf{f}) = P[f(1)]\,P[f(Q)\,|\,f(1)]\,P[f(2)\,|\,f(Q),\ f(1)]\ \ldots$$
$$\ldots P[f(q)\,|\,f(Q),\ f(Q-1),\ \ldots,\ f(q+1),\ f(q-1),\ \ldots$$
$$\ldots,\ f(2), f(1)]. \tag{7.6.11}$$

Повторяя рассуждения, которые привели к соотношению (7.6.7), энтропию источника изображений можно выразить приближенным равенством

$$H(\mathbf{f}) \approx QH\,[f(q)\,|\,f(Q),\ f(Q-1),\ \ldots,\ f(q+1),\ f(q-1),\ \ldots,$$
$$f(2),\ f(1)]. \tag{7.6.12}$$

Полученное «двустороннее» выражение для условной энтропии можно в свою очередь аппроксимировать, учитывая только ближайшие элементы столбца:

$$H(\mathbf{f}) \approx QH\,[f(q)\,|\,f(q+k),\ \ldots,\ f(q+1),\ f(q-1),\ \ldots,$$
$$f(q-k)]. \tag{7.6.13}$$

Развивая эту идею, можно включить в выражение для оценки энтропии все элементы вектора, которые имеют достаточно сильные статистические связи с $f(q)$. Во многих случаях такие элементы оказываются геометрически ближайшими к исследуемому. Поэтому можно принять следующую оценку энтропии:

$$H(\mathbf{f}) \approx QH\,[f(q)\,|\,f(q-1),\ f(q+1),\ f(q-N),$$
$$f(q+N)], \tag{7.6.14}$$

где N — длина столбца, а энтропия элемента $f(q)$ определяется значениями ближайших четырех элементов, находящихся сверху, снизу, справа и слева от него [1].

Прейс [44] вычислял энтропию двухградационных факсимильных документов, используя несколько предшествующих элементов развертки. К сожалению, из-за чрезмерного объема необходимых вычислений трудно оценить энтропию многоградационных изображений даже с использованием упрощенной формулы (7.6.14). Для вычислений по этой формуле необходимо получить распределение частот пятого порядка, причем число возможных значений каждого аргумента равно числу уровней квантования яркости. Приходится делать печальный вывод, что вычисление энтропии в принципе позволяет оценивать «содержательность» изображения, однако для многоградационных изображений такие вычисления практически невыполнимы [2].

7.7. СКОРОСТЬ СОЗДАНИЯ ИНФОРМАЦИИ ИСТОЧНИКОМ ИЗОБРАЖЕНИЙ

Скорость создания информации некоторым источником определяет минимальную величину пропускной способности канала, которая необходима для передачи сообщения с заданным допустимым уровнем искажений в принятом сообщении относительно переданного [45, 46]. Был предпринят ряд попыток [2, 47—50] при-

Рис. 7.7.1. Блок-схема системы передачи информации.

способить теорию информации к задачам передачи изображений с тем, чтобы определить предельные возможности систем кодирования изображений. В данном разделе приведены основные положения этой теории, сформулированные применительно к изображениям на основе обзорной статьи Дэвиссона [50].

На рис. 7.7.1 приведена упрощенная блок-схема системы передачи изображений. Источник создает последовательность из Q элементов изображения, каждый из которых квантуется на J уровней. Эта последовательность образует вектор $\mathbf{f}$ размера $Q \times 1$. В кодере каждому из $T = J^Q$ возможных сочетаний яркостей $\mathbf{f}_i$, где $i = 1, 2, \ldots, T$, ставится в соответствие кодовая комбинация.

[1] Вывод выражения для энтропии изображения, когда все его элементы поступают одновременно, не вполне корректен. Оценка (7.6.14) дает заниженное значение энтропии. — *Прим. ред.*

[2] Можно получить оценки энтропии изображения, используя его вероятностные модели. — *Прим. ред.*

После декодирования восстанавливается сочетание яркостей $\hat{\mathbf{f}}_k$. Характеристики системы передачи изображений можно описать с помощью условной вероятности $P[\hat{\mathbf{f}}_k \mid \mathbf{f}_i]$ вектора $\hat{\mathbf{f}}_k$ на выходе при условии, что кодированию подвергался вектор $\mathbf{f}_i$. Если кодер и декодер работают без ошибок, то входной и выходной векторы изображения (при отсутствии ошибок в канале) будут одинаковыми. Условная вероятность $P[\hat{\mathbf{f}}_k \mid \mathbf{f}_i]$ описывает работу системы передачи изображений при наличии искажений. На основании этой условной вероятности и распределения априорных вероятностей находим безусловное распределение вероятностей восстановленных векторов

$$P(\hat{\mathbf{f}}_k) = \sum_{i=1}^{T} P(\hat{\mathbf{f}}_k \mid \mathbf{f}_i)\, P(\mathbf{f}_i). \tag{7.7.1}$$

Требования к пропускной способности канала определяются количеством взаимной информации, по определению равной

$$I_Q(\mathbf{f}, \hat{\mathbf{f}}) = \sum_{i=1}^{T} \sum_{k=1}^{T} P(\mathbf{f}_i)\, P(\hat{\mathbf{f}}_k \mid \mathbf{f}_i) \log [P(\hat{\mathbf{f}}_k \mid \mathbf{f}_i)/P(\hat{\mathbf{f}}_k)]. \tag{7.7.2}$$

При безошибочном кодировании это выражение упрощается:

$$I_Q(\mathbf{f}, \hat{\mathbf{f}}) = -\sum_{i=1}^{T} P(\mathbf{f}_i) \log [P(\mathbf{f}_i)] \equiv H(\mathbf{f}), \tag{7.7.3}$$

т. е. количество взаимной информации оказывается равным энтропии источника. Если в процессе кодирования вносятся искажения, то восстановленная последовательность $\hat{f}_i$ содержит неполную информацию о состоянии f_i и требования к пропускной способности канала будут уменьшены.

Допустим, что функция $D(\mathbf{f}, \hat{\mathbf{f}})$ представляет собой некоторую меру искажений воспроизведенного изображения. Тогда для вектора из Q элементов средняя величина искажений в расчете на один элемент будет определяться равенством

$$D = \frac{1}{Q} E\{D(\mathbf{f}, \hat{\mathbf{f}})\} = \frac{1}{Q} \sum_{i=1}^{T} \sum_{k=1}^{T} D(\mathbf{f}_i, \hat{\mathbf{f}}_k)\, P(\mathbf{f}_i)\, P(\hat{\mathbf{f}}_k \mid \mathbf{f}_i). \tag{7.7.4}$$

Определим для этого вектора скорость создания информации в расчете на один элемент как

$$R_Q(D_*) = \inf \left[\frac{1}{Q} I_Q(\mathbf{f}, \hat{\mathbf{f}})\right] \tag{7.7.5}$$

при $D < D_*$ [1]. В принципе $R_Q(D_*)$ и есть минимальная пропускная способность канала, необходимая для передачи информации,

[1] Символ inf [·] обозначает инфинум аргумента.

создаваемой источником, когда искажения в среднем не должны превышать некоторой максимальной величины D_*. Скорость создания информации источником $R(D_*)$ можно найти, увеличивая длину вектора до бесконечности:

$$R(D_*) = \lim_{Q\to\infty} [R_Q(D_*)]. \tag{7.7.6}$$

Как правило, отыскать минимум количества взаимной информации при условии, что средняя величина искажений не должна превышать заданного предела D_*, сложно как аналитическими, так и численными методами. Получено несколько решений для каналов связи, используемых на практике. Одно из них относится к источнику с гауссовым распределением вероятностей при оценке искажений среднеквадратической мерой. Такое решение нельзя непосредственно применить к задаче о кодировании яркостей элементов, поскольку яркость — неотрицательная величина. Кроме того, среднеквадратическая мера искажений может оказаться неподходящей. Однако решение, полученное для гауссова источника и среднеквадратической меры искажений, позволяет указать предельные возможности системы кодирования для любых источников с заданными вторыми моментами. Кроме того, это решение непосредственно переносится на задачу о кодировании с преобразованием. Поэтому ниже будут рассмотрены свойства скорости создания информации для случая гауссова источника и среднеквадратической меры искажений.

Рассмотрим вектор $\mathbf{f}$, образованный Q независимыми гауссовыми случайными величинами с нулевыми средними и известной дисперсией σ^2. Среднеквадратическая ошибка определяется формулой

$$D(\mathbf{f}, \hat{\mathbf{f}}) = \frac{1}{Q} \sum_{q=1}^{Q} E\{[f(q) - \tilde{\mathbf{f}}(q)]^2\}. \tag{7.7.7}$$

Было найдено [45], что скорость создания информации

$$R(D_*) = R_Q(D_*) = \begin{cases} {}^1/_2 \log_2(\sigma^2/D_*) \text{ при } \sigma^2 > D_*, & (7.7.8\text{а}) \\ 0 \text{ при } \sigma^2 \leqslant D_*. & (7.7.8\text{б}) \end{cases}$$

Таким образом, скорость создания информации равна половине логарифма отношения мощности сигнала к мощности искажений, если это отношение превышает единицу, и нулю в противном случае. Если элементы последовательности, создаваемой гауссовым источником, коррелированы и ковариационная матрица $\mathbf{K}_f$ известна, то скорость создания информации равна [50]

$$R_Q(D_*) = \frac{1}{Q} \sum_{q=1}^{Q} \frac{1}{2} \log_2 \left\{ \frac{\lambda(q)}{\min[\lambda(q), D_{**}]} \right\}, \tag{7.7.9}$$

где $\lambda(q)$ — q-е собственное значение матрицы $\mathbf{K}_f$, а D_{**} выбрано так, что

$$D_* = \frac{1}{Q} \sum_{q=1}^{Q} \min[\lambda(q),\ D_{**}]. \tag{7.7.10}$$

При обработке изображений представляет интерес частный случай двумерного разделимого марковского источника, когда все элементы имеют одинаковые дисперсии, равные σ^2, а коэффициенты

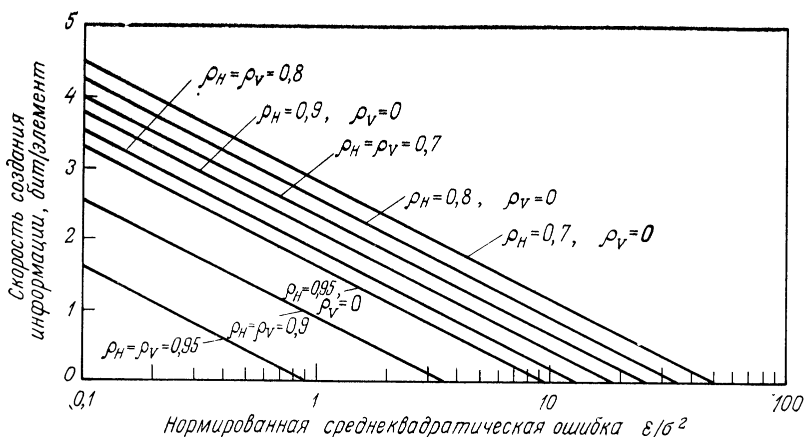


Рис. 7.7.2. Зависимость скорости создания информации от величины искажений при одно- и двумерном кодировании изображений — реализаций марковского поля.

корреляции вдоль строк и столбцов равны соответственно ρ_R и ρ_C. Если предположить, что степень искажений невелика, то скорость создания информации для однородного случая равна [50]

$$R(D_*) = \frac{1}{2} \log_2 \left[\frac{\sigma^2 (1 - \rho_D^2)}{D_*} \right], \tag{7.7.11}$$

а в двумерном случае

$$R(D_*) = \frac{1}{2} \log_2 \left[\frac{\sigma^2 (1 - \rho_R^2)(1 - \rho_C^2)}{D_*} \right]. \tag{7.7.12}$$

На рис. 7.7.2 приведены графики зависимости скорости создания информации от величины искажений при различных значениях коэффициентов корреляции. В гл. 24 проведено сравнение характеристик некоторых систем кодирования изображений с предельными характеристиками.

7.8. ДЕШИФРИРУЕМОСТЬ ИЗОБРАЖЕНИЙ

Определение количественной меры дешифрируемости изображения, т. е. степени его пригодности для анализа наблюдателем, является гораздо более сложной задачей, чем измерение верности воспроизведения изображений. Факторы, влияющие на дешифрируемость, глубоко связаны со свойствами человеческого зрения. Способность человека-наблюдателя извлекать из изображений информацию зависит, очевидно, от физических свойств зрительной системы, которые определяются, в частности, оптическими свойствами глаза, нелинейной реакцией фоторецепторов на яркость и наличием латерального торможения. Однако алгоритмы обработки данных, поступающих от глаза в мозг, играют важную и, по-видимому, решающую роль в процессе извлечения человеком визуальной информации из изображения. Поскольку механизмы зрительного восприятия пока изучены слабо, то в настоящее время попытки создать критерий дешифрируемости изображений, основанный на модели зрительной системы человека, представляются безнадежными. Следует скорее остановиться на каких-то частных критериях, которые достаточно хорошо согласовывались бы с результатами субъективных оценок.

В большинстве экспериментов по экспертной оценке дешифрируемости изображений используются простые штриховые фигуры, как, например, оптометрические таблицы, составленные из подобранных случайным образом букв разного размера. На рис. 7.8.1 представлены четыре контрольные таблицы, применявшиеся в экспериментах по распознаванию объектов [51]. Таблицы Ландольта [52, 53] составлены из строк буквы С, которая может находиться в любом из четырех (или восьми) возможных положений, причем размеры букв от строки к строке уменьшаются. Цифровая таблица составлена из «криволинейных» цифр 3, 5, 6, 8 и 9, ориентированных случайным образом. В таблице из звездочек в каждом символе отсутствует какой-то один из шести лучей. Наблюдатель должен распознать символы во всех таблицах; результаты эксперимента выражаются в виде зависимости вероятности правильного ответа от размера символа. Подобные эксперименты на дискретизованных вариантах четырех таблиц, показанных на рис. 7.8.1, проводил Бернард [51]. Результаты приведены на рис. 7.8.2. В этих экспериментах были дискретизованы изображения на диапозитивах, затем к отсчетам прибавляли гауссов шум и результаты записывались на фотопленке в виде диапозитива. Как и ожидалось, с уменьшением размеров символов вероятность распознавания монотонно убывает. Бернард заметил также, что в экспериментах с цифрами наблюдается больший разброс, чем в опытах со звездочками или с таблицами Ландольта. Кроме того, оказалось, что вероятность правильного различения цифр зависит от их ориентации.

Так, при распознавании повернутой цифры «3» ошибки делали чаще, чем при распознавании такой же цифры, стоящей в нормальном положении. Можно предполагать, что в экспериментах со

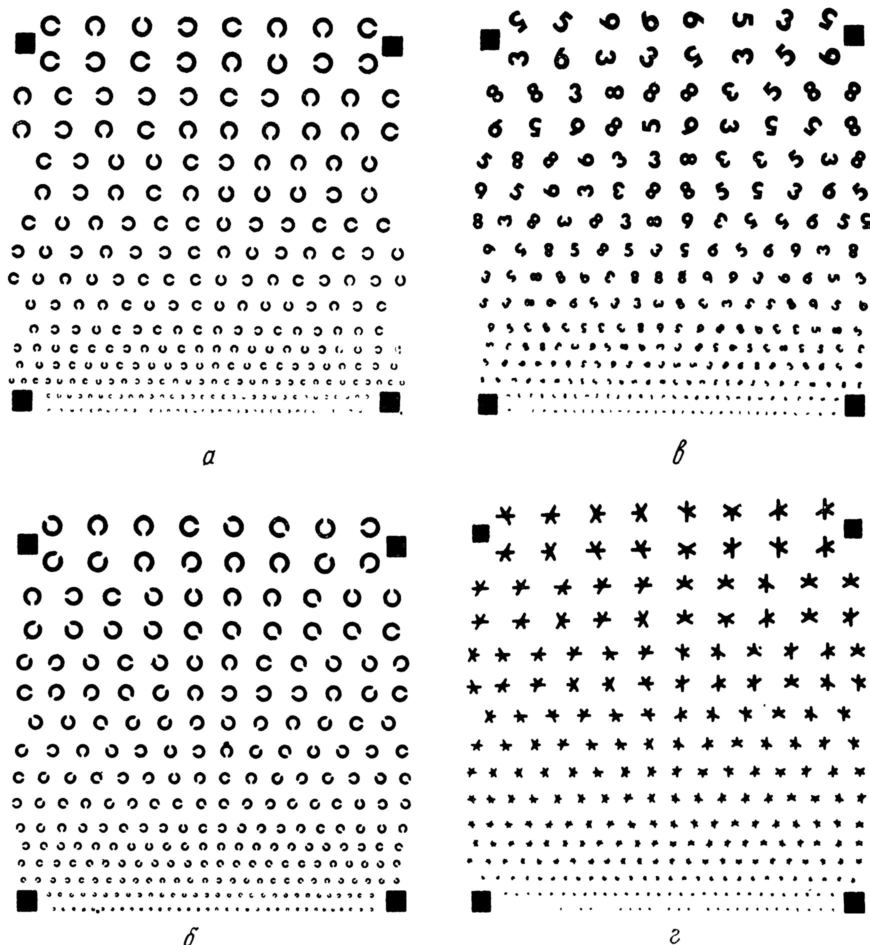


Рис. 7.8.1. Испытательные таблицы для экспериментов по распознаванию символов [51].

а — кольца Ландольта с четырьмя позициями; *б* — кольца Ландольта с восемью позициями; *в* — цифры; *г* — звездочки

звездочками и таблицами Ландольда такой зависимости не наблюдалось потому, что у этих символов нет «стандартной» ориентации. Бернард [51] для предсказания результатов эксперимента по различению человеком символов в таблицах Ландольта применял

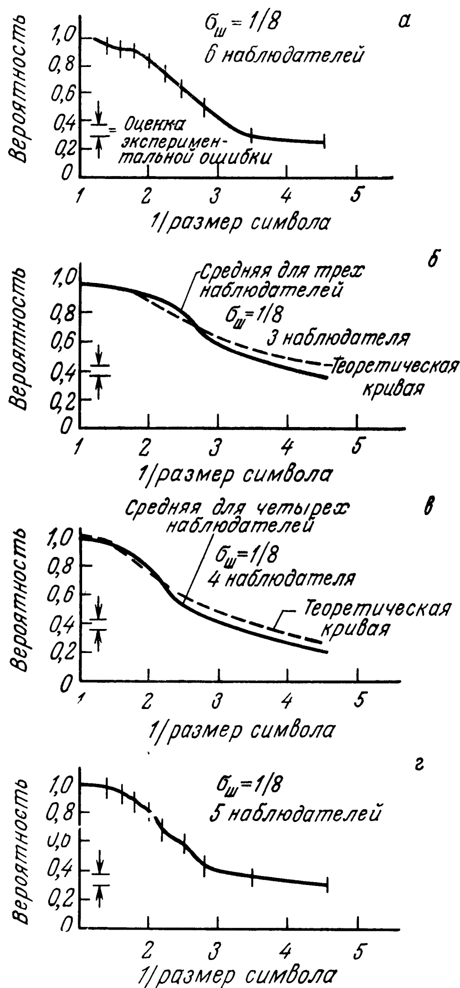


Рис. 7.8.2. Результаты экспериментов по распознаванию символов [51].

а — таблица цифр; *б* — таблица с кольцами Ландольта (четыре позиции); *в* — таблица с кольцами Ландольта (восемь позиций); *г* — таблица звездочек

теорию согласованной фильтрации [1]. Вероятность правильного распознавания M ортогональных сигналов (M ориентаций символа) в присутствии аддитивного гауссова шума аппроксимировалась [54] выражением

$$P = (1/\sqrt{\pi}) \int_{-\infty}^{\infty} \exp\{-z^2\} [{}^1/_2 + {}^1/_2 \operatorname{erf}\{\sqrt{{}^1/_2 (S/N)} + z\}]^{M-1} dz, \tag{7.8.1}$$

[1] Анализ свойств двумерных согласованных фильтров представлен в гл. 19.

где S/N — отношение мощностей сигнала и шума. Как показано на рис. 7.8.2, *б*, *в*, теоретические и экспериментальные результаты согласуются довольно хорошо.

Был проведен ряд опытов по оценке разборчивости цифробуквенных символов, переданных с помощью фототелеграфных аппаратов. Арпс и др. [55] обнаружили, что для достижения 97,5 %-ной вероятности правильного распознавания необходимо, чтобы разрешение при дискретизации символа было порядка одной десятой наибольшего размера символа. Эрдман и Нил [56] изучали, как изменяется разборчивость слов в зависимости от разборчивости букв, связности текста и других семантических факторов.

Проводились также исследования с целью определения «рабочих характеристик» человека при визуальном анализе многоградационных изображений [57]. Однако исследования в этом направлении все еще находятся на начальной стадии своего развития.

ЛИТЕРАТУРА

1. Pearson D. E., Methods for Scaling Television Picture Quality, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972.
2. Mannos J. L., Sakrison D. J., The Effects of a Visual Fidelity Criterion on the Encoding of Images, *IEEE Trans. Inf. Theory*, **IT-20**, 4, 525—536 (July 1974).
3. Mertz P., Fowler A. D., Christoper H. N., Quality Rating of Television Images, *Proc. IRE*, **38**, 11, 1269—1283 (November 1950).
4. Fredenhall G. L., Behrend W. L., Picture Quality-Procedures for Evaluating Subjective Effects of Interference, *Proc. IRE*, **48**, 6, 1030—1034 (June 1960).
5. Dean C. E., Measurements of the Subjective Effects of Interference in Television Reception, *Proc. IRE*, **48**, 6, 1035—1050 (June 1960).
6. Prosser R. D, Allnatt J. W., Lewis N. W., Quality Grading of Impaired Television Pictures, *Proc. IEE*, **111**, 3, 491—502 (March 1964).
7. Weaver L. E., The Quality Rating of Color Television Pictures, *J. Soc. Motion Picture Television Engineers*, **77**, 6, 610—612 (June 1968).
8. Lewis N. W., Waveform Responses of Television Links, *Proc. IEE*, **101**, part III, 258—270 (1954).
9. Fisher J. F., Generation of NTSC Color Signals, *Proc. IRE*, **41**, 3, 338—343 (March 1953).
10. Luther A. C., Methods of Verifying Adherence to the NTSC Color Signal Specifications, *Proc. IRE*, **42**, 1, 235—240 (January 1954).
11. Fink D. G., Ed., Television Engineering Handbook, McGraw-Hill, New York, 1957.
12. Cornsweet T. N., Visual Perception, Academis Press, New York, 1970.
13. Priest I. G., Gibson K. S., McNicholas H. J., An Examination of the Munsell Color System, I. Spectral and Total Reflection and the Muncell Scale of Value, U. S. National Bureau of Standards, Technical Paper 167, 1920.
14. Ladd J. H, Pinney J. E., Empirical Relationships with the Munsell Value Scale, *Proc. IRE* (Correspondence), **43**, 9, 1137 (1955).
15. Foss C. E., Nickerson D., Granville W. C., Analysis of the Ostwald Color System, *J. Opt. Soc. Am.*, **34**, 7, 361—381 (July 1944).
16. Wyszecki G. W., Stiles W. S., Color Science, Wiley, New York, 1967.

17. Judd D. B., Hue, Saturation, and Lightness of Surface Colors with Chromatic Illumination, *J. Opt. Soc. Am.*, **30**, 1, 2—32 (January 1940).
18. Wright W. D., The Sensitivity of the Eye to Small Colour Differences, *Proc. Phys. Soc* (*London*), **53**, 93 (1941).
19. Muth E. J., Persels C. G., Computation of Geodesics in Color 3-Space by Dynamic Programming, Proceedings 4th Hawaii Conference on System Sciences, January 1971, pp. 155—157.
20. Jain A. K., Color Distance and Geodesics in Color 3 Space, *J. Opt. Soc. Am.*, **62**, 11, 1287—1290 (November 1972).
21. Judd D. B., A Maxwell Triangle Yielding Uniform Chromaticity Scales, *J. Opt. Soc. Am.*, **25**, 1, 24—35 (January 1935).
22. MacAdam D. L., Projective Transformations of ICI Color Specifications, *J. Opt. Soc. Am.*, **27**, 9, 294—299 (August 1937).
23. MacAdam D L., Visual Sensitivities to Color Differences in Daylight, *J. Opt Soc. Am.*, **32**, 5, 247—274 (May 1942).
24. Farnsworth D., A Temporal Factor in Colour Discrimination, Visual Problems in Colour, II. National Phys. Lab. Symposium, No. 8, London, Her Majesty's Stationary Office, 1958, p. 429.
25. Wyszecki G., Proposal for a New Color-Difference Formula, *J. Opt. Soc. Am.*, **53**, 11, 1318—1319 (November 1963).
26. Hunt R. W. G., The Reproduction of Colour, Wiley, New York, 1967.
27. Barnard T. W., A Literature Survey on Image Quality Evaluation, The Perkin Elmer Corporation, Engineering Report ER-177, June 1971.
28. Frieden B. R., Image Evaluation by Use of the Sampling Theorem, *J. Opt. Soc. Am.*, **56**, 10, 1355—1362 (October 1966).
29. Linfoot E. H., Quality Evaluations of Optical Systems, *Optica Acta*, **5**, 1—2, 1—14 (March — June 1958).
30. O'Neill E. L., Introduction to Statistical Optics, Addison-Wesley, Reading, Mass., 1963. [Имеется перевод: О'Нейл Э., Введение в статистическую оптику. — М.: Мир, 1966.]
31. Harris J. L., Sr., Image Evaluation and Restoration, *J. Opt. Soc. Am* , **56**, 5, 569—574 (May 1966).
32. Harris J. L., Sr., Resolving Power and Decision Theory, *J. Opt. Soc. Am.*, **54**, 5, 606—611 (May 1964).
33. Born M., Wolf E., Principles of Optics, 4th ed., Pergamon, New York, 1970 [Имеется перевод: Борн М., Вольф Э., Основы оптики. — М.: Наука, 1973.]
34. Boyce B., A Summary Measure of Image Quality, Proceedings of a Symposium on Current Mathematical Problems in Image Science, Naval Postgraduate School, Monterey, California, November 1976.
35. Schade O. H., Modern Image Evaluation and Television, *Appl. Opt* , **3**, 1, 17—23 (January 1964).
36 Andrews H. C., Computer Techniques in Image Processing, Academic Press, New York, 1970. [Имеется перевод: Эндрюс Г., Применение вычислительных машин для обработки изображений. — М.: Энергия, 1977.]
37. Wilder W. C., Subjectively Relevant Error Criteria for Pictorial Data Processing, Purdue University, School of Electrical Engineering, Report TR-EE 72-34, December 1972.
38 Pratt W. K., Spatial Transform Coding of Color Images, *IEEE Trans. Commun Tech* , **COM-19**, 6, 980—992 (December 1971).
39. Shannon C. E., The Mathematical Theory of Communication, University of Illinois Press, Urbana, Illinois, 1949. (Orig. Pub., *Bell Syst. Tech. J* , 1948.) [Имеется перевод: Шеннон К., Работы по теории информации и кибернетике. — М.: ИЛ, 1963, с. 243—332.]
40. Abramson N , Information Theory and Coding, McGraw-Hill, New York, 1963.
41. Ash R. B., Information Theory and Reliable Communication, Wiley, New York, 1968.

42. Gallagher R. G., Information Theory, Wiley-Interscience, New York, 1965. [Имеется перевод: Галлагер Р., Теория информации и надежная связь. — М.: Советское радио, 1974.]
43. Schreiber W. F., The Measurement of Third Order Probability Distributions of Television Signals, *IRE Trans. Inf. Theory*, **IT-2**, 3, 94—105 (September 1956).
44. Preuss D., Comparison of Two-Dimensional Facsimile Coding Schemes, International Communications Conference, 1975 Conference Record, Vol. 1, June 1975, pp. 7.12—7.16.
45. Berger T., Rate Distortion Theory: A Mathematical Basis for Data Compression, Prentice-Hall, Englewood Cliffs, N. J., 1971.
46. Andrews H. C., Bibliography on Rate Distortion Theory, *IEEE Trans. Inf. Theory* (*Correspondence*), **IT-17**, 198—199 (March 1971).
47. Hayes J. F., Habibi A., Wintz P. A., Rate Distortion Function for a Gaussian Source Model of Images, *IEEE Trans. Inf. Theory* (*Correspondence*), **IT-16**, 507—508 (July 1970).
48. Sakrison D. J., Algazi V. R., Comparison of Line-by-Line and Two-Dimensional Encoding of Random Images, *IEEE Trans. Inf. Theory*, **IT-17**, 386—398 (July 1971).
49. Tasto M., Wintz, P. A., A Bound on the Rate-Distortion Function and Application of Images, *IEEE Trans. Inf. Theory*, **IT-18**, 150—159 (January 1972).
50. Davisson L. D., Rate Distortion Theory and Applications, *Proc. IEEE*, **60**, 7, 800—808 (July 1971). [Имеется перевод: Дэвиссон. Скорость создания сообщений. Теория и применение. — ТИИЭР, 1972, т. 60, № 7.]
51. Barnard T. W., An Image Evaluation Method, A Symposium on Sampled Images, Perkin-Elmer Corporation, Norwalk, Connecticut, 1971.
52. Landolt E., Tableau d'Optotypes pour la Determination de l'Acuite Visuel, Society François d'Ophthol, 1889, p. 157.
53. Cowan A., Test Cards for Determination of Visual Acuity, Arch. Ophthol, **57**, 283—295 (1928).
54. Schwartz M., Bennett W. R., Stein S., Communication Systems and Techniques, McGraw-Hill, New York, 1966, p. 85.
55. Arps R. B. et al., Character Legibility Versus Resolution in Image Processing of Printed Matter, *IEEE Trans. Man-Machine Systems*, **MMS-10**, 3, 66—71 (September 1969).
56. Erdmann R. L., Neal A. S., Word Legibility as a Function of Letter Legibility with Word Size, Word Familiarity, and Resolution as Parameters, *J. Appl. Psych.*, **52**, 5, 403—409 (1968).
57. Snyder H. L., Image Quality and Observer Performance, in: Perception of Displayed Information, Biberman L. M., Ed., Plenum Press, New York, 1973.

Часть 3

Дискретная линейная двумерная обработка

В третьей части книги с единых позиций анализируются дискретные линейные операции, выполняемые при двумерной обработке. Даны определения обобщенных линейных операторов и операторов псевдообращения. Введено несколько видов дискретных двумерных операторов суперпозиции и свертки и установлены соотношения между ними. Представлены двумерные преобразования, такие, как преобразования Фурье, Адамара, Карунена—Лоэва, и описаны быстрые алгоритмы вычисления. Рассмотрен вопрос о применении двумерных преобразований как эффективного средства выполнения свертки. В заключение описаны рекуррентные методы линейной обработки изображений.

Глава 8

ЛИНЕЙНЫЕ ОПЕРАТОРЫ

Многие операции, выполняемые при обработке изображений, по своей природе являются линейными. В число таких операций входят суперпозиция, свертка, унитарные преобразования и дискретная линейная фильтрация.

8.1. ОБОБЩЕННЫЙ ЛИНЕЙНЫЙ ОПЕРАТОР

Рассмотрим массив $F(n_1, n_2)$ из $N_1 \times N_2$ элементов, представляющий исходное (входное) изображение. При воздействии на него обобщенным линейным оператором получается массив из $M_1 \times M_2$ элементов, описывающий преобразованное (выходное) изображение

$$P(m_1, m_2) = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} F(n_1, n_2) O(n_1, n_2; m_1, m_2), \qquad (8.1.1)$$

где ядро оператора $O(n_1, n_2; m_1, m_2)$ представляет собой набор весовых множителей, зависящих в общем случае от координат элементов как входного, так и выходного изображений.

При анализе линейных операций по обработке изображений удобно пользоваться векторными представлениями, описанными в гл. 5 [1]. Поэтому будем полагать, что входной массив $F(n_1, n_2)$ представлен или в виде матрицы $\mathbf{F}$, или в виде вектора $\mathbf{f}$, полученного разверткой матрицы $\mathbf{F}$ по столбцам. Аналогичным образом допустим, что выходной массив $P(m_1, m_2)$ может быть представлен либо матрицей $\mathbf{P}$, либо в виде ее развертки по столбцам, т. е. вектором $\mathbf{p}$. Для упрощения обозначений ниже будет принято, что матрицы, представляющие входное и выходное изображения, квадратные с размерами $N_1 = N_2 = N$ и $M_1 = M_2 = M$ соответственно. Допустим теперь, что символ $\mathbf{T}$ обозначает матрицу размера $M^2 \times N^2$, с помощью которой вектор исходного изображения $\mathbf{f}$ размера $N^2 \times 1$ линейно преобразуется в вектор

$$\mathbf{p} = \mathbf{T}\mathbf{f} \qquad (8.1.2)$$

выходного изображения размера $M^2 \times 1$. Матрицу $\mathbf{T}$ можно разделить на блоки — матрицы $\mathbf{T}_{mn}$ размера $M \times N$ (число которых так-

же равно $M \times N$) — и представить ее в следующем виде:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \ldots & \mathbf{T}_{1N} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \ldots & \mathbf{T}_{2N} \\ \cdot & \cdot & \cdot & \cdot \\ \mathbf{T}_{M1} & \mathbf{T}_{M2} & \ldots & \mathbf{T}_{MN} \end{bmatrix}. \tag{8.1.3}$$

В согласии с соотношением (5.3.3) вектор выходного изображения **p** можно выразить через матрицу входного изображения **F**:

$$\mathbf{p} = \sum_{n=1}^{N} \mathbf{T}\mathbf{N}_n\mathbf{F}\mathbf{v}_n. \tag{8.1.4}$$

Кроме того, с помощью равенства (5.3.4) матрицу **P** выходного изображения можно выразить через вектор **p** того же изображения:

$$\mathbf{P} = \sum_{m=1}^{M} \mathbf{M}_m^T\mathbf{p}\mathbf{u}_m^T. \tag{8.1.5}$$

Из этих формул получим выражение, связывающее между собой входную и выходную матрицы:

$$\mathbf{P} = \sum_{m=1}^{M} \sum_{n=1}^{N} (\mathbf{M}_m^T\mathbf{T}\mathbf{N}_n)\,\mathbf{F}\,(\mathbf{v}_n\mathbf{u}_m^T). \tag{8.1.6}$$

Заметим, что операторы $\mathbf{M}_m$ и $\mathbf{N}_n$ просто выделяют из матрицы **T** блок $\mathbf{T}_{mn}$. Следовательно,

$$\mathbf{P} = \sum_{m=1}^{M} \sum_{n=1}^{N} \mathbf{T}_{mn}\mathbf{F}\,(\mathbf{v}_n\mathbf{u}_m^T). \tag{8.1.7}$$

Пусть линейное преобразование является разделимым, т. е. матрицу **T** можно представить в виде прямого произведения

$$\mathbf{T} = \mathbf{T}_C \otimes \mathbf{T}_R, \tag{8.1.8}$$

где $\mathbf{T}_C$ и $\mathbf{T}_R$ — операторы преобразования столбцов и строк матрицы изображения **F**. В этом случае

$$\mathbf{T}_{mn} = \mathbf{T}_R(m, n)\,\mathbf{T}_C. \tag{8.1.9}$$

Следовательно,

$$\mathbf{P} = \mathbf{T}_C\mathbf{F} \sum_{m=1}^{M} \sum_{n=1}^{N} T_R(m, n)\,\mathbf{v}_n\mathbf{u}_m^T = \mathbf{T}_C\mathbf{F}\mathbf{T}_R^T. \tag{8.1.10}$$

Таким образом, матрицу выходного изображения **P** можно получить путем последовательной обработки матрицы **F** по строкам и столбцам.

При обработке изображений во многих случаях оказывается, что оператор линейного преобразования **T** имеет специфическую структуру, позволяющую упростить вычислительные операции. Ниже перечислены важные частные случаи, проиллюстрированные

на рис. 8.1.1, когда размеры входного и выходного изображений выбраны одинаковыми, т. е. $M = N$.

а) При обработке матрицы **F** только по столбцам

$$\mathbf{T} = \mathrm{diag}\,[\mathbf{T}_{C1}, \mathbf{T}_{C2}, \ldots, \mathbf{T}_{CN}], \tag{8.1.11}$$

где $\mathbf{T}_{Cj}$ — матрица преобразования для j-го столбца.

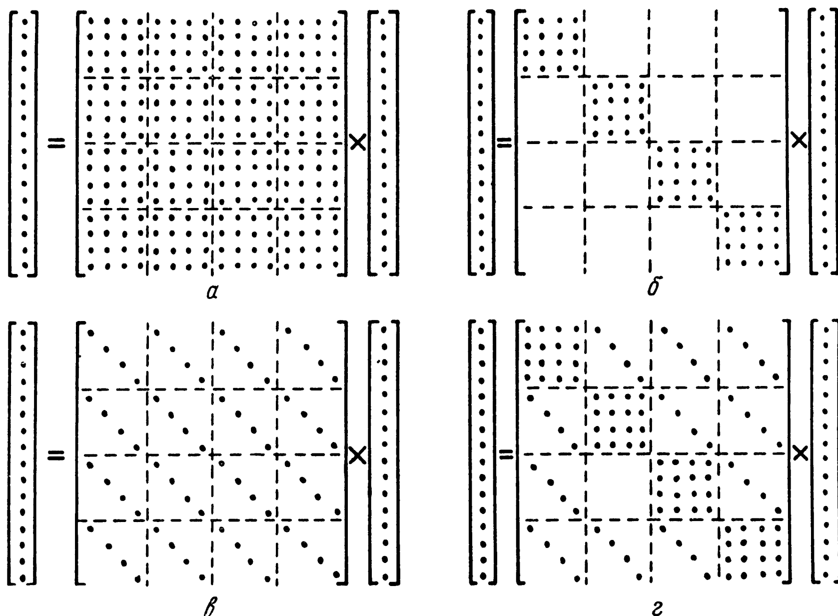


Рис. 8.1.1. Структура матриц линейного оператора.

а — общий случай; *б* — обработка только по столбцам; *в* — обработка только по строкам; *г* — обработка только по строкам и столбцам.

б) При одинаковой обработке каждого столбца матрицы **F**

$$\mathbf{T} = \mathrm{diag}\,[\mathbf{T}_C, \mathbf{T}_C, \ldots, \mathbf{T}_C] = \mathbf{T}_C \otimes \mathbf{I}_N. \tag{8.1.12}$$

в) При обработке матрицы **F** только по строкам

$$\mathbf{T}_{mn} = \mathrm{diag}\,[\mathbf{T}_{R1}(m, n), \ldots, \mathbf{T}_{RN}(m, n)], \tag{8.1.13}$$

где $\mathbf{T}_{Rj}$ — матрица преобразования для j-й строки.

г) При одинаковой обработке каждой строки матрицы **F**

$$\mathbf{T}_{mn} = \mathrm{diag}\,[T_R(m, n), T_R(m, n), \ldots, T_R(m, n)], \tag{8.1.14а}$$

$$\mathbf{T} = \mathbf{I}_N \otimes \mathbf{T}_R. \tag{8.1.14б}$$

д) При одинаковой обработке столбцов и одинаковой обработке строк матрицы **F**

$$\mathbf{T} = \mathbf{T}_C \otimes \mathbf{I}_N + \mathbf{I}_N \otimes \mathbf{T}_R. \tag{8.1.15}$$

Число арифметических операций, выполняемых в каждом из этих случаев, указано в табл. 8.1.1.

Из соотношения (8.1.10) видно, что если двумерное линейное преобразование имеет разделимую матрицу, то его можно выполнить путем последовательной одномерной обработки строк и столбцов массива отсчетов. Как следует из табл. 8.1.1, для таких преобразований удается существенно сократить число необходимых вычислительных операций: в общем случае при вычислении по формуле (8.1.2) требуется M^2N^2 операций, но если можно воспользоваться формулой (8.1.10), достаточно $MN^2 + M^2N$ операций. Более того, в этом случае матрицу **F** можно хранить в запоминающем устройстве (ЗУ) с последовательным доступом, например на диске или на барабане, и считывать строка за строкой, т. е. отпадает необходимость хранения матрицы **F** в более дорогостоящем ЗУ с произвольным доступом. Необходимо, однако, транспонировать результаты преобразований по столбцам с тем, чтобы выполнить построчные преобразования. В работах [2, 3] описаны алгоритмы транспонирования матриц, записанных в ЗУ с последовательным доступом.

Таблица 8 1.1

Число арифметических операций при линейном преобразовании

Случай	Число умножений и сложений
Общий	N^4
Обработка по столбцам	N^3
Обработка по строкам	N^3
Обработка по строкам и столбцам	$2N^3 - N^2$
Обработка при разделимой матрице	$2N^3$

8.2. СТАТИСТИЧЕСКОЕ ОПИСАНИЕ ИЗОБРАЖЕНИЙ ПРИ ЛИНЕЙНОМ ПРЕОБРАЗОВАНИИ

Если массив отсчетов входного изображения является реализацией случайного процесса с известными первым и вторым моментами, то для заданного линейного преобразования можно найти первый и второй моменты также и для массива отсчетов выходного изображения. Среднее значение выходного массива равно

$$E\{P(m_1, m_2)\} = E\left\{\sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} F(n_1, n_2)\, O(n_1, n_2; m_1, m_2)\right\}. \tag{8 2.1a}$$

Однако оператор усреднения линеен, и поэтому

$$E\{P(m_1, m_2)\} = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} E\{F(n_1, n_2)\} O(n_1, n_2; m_1, m_2). \tag{8.2.1б}$$

Корреляционная функция массива отсчетов выходного изображения имеет вид

$$R_P(m_1, m_2; m_3, m_4) = E\{P(m_1, m_2) P^*(m_3, m_4)\} =$$

$$= E\left\{\left[\sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} F(n_1, n_2) O(n_1, n_2; m_1, m_2) \sum_{n_3=1}^{N_1} \sum_{n_4=1}^{N_2} F^*(n_3, n_4) \times \right.\right.$$

$$\left.\left. \times O^*(n_3, n_4; m_3, m_4)\right]\right\}. \tag{8.2.2а}$$

После перемножения рядов и вычисления среднего получается выражение

$$R_P(m_1, m_2; m_3, m_4) =$$

$$= \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \sum_{n_3=1}^{N_1} \sum_{n_4=1}^{N_2} R_F(n_1, n_2; n_3, n_4) O(n_1, n_2; m_1, m_2) \times$$

$$\times O^*(n_3, n_4; m_3, m_4), \tag{8.2.2б}$$

где $R_F(n_1, n_2; n_3, n_4)$ — корреляционная функция массива отсчетов исходного изображения. Аналогично определяется и ковариационная функция обработанного изображения

$$K_P(m_1, m_2; m_3, m_4) =$$

$$= \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \sum_{n_3=1}^{N_1} \sum_{n_4=1}^{N_2} K_F(n_1, n_2; n_3, n_4) O(n_1, n_2; m_1, m_2) \times$$

$$\times O^*(n_3, n_4; m_3, m_4). \tag{8.2.3}$$

Если массивы отсчетов входного и выходного изображений представлены в векторной форме, то выражения для их моментов можно записать более компактно. Среднее значение выходного вектора $\mathbf{p}$

$$\eta_{\mathbf{p}} = E\{\mathbf{p}\} = E\{\mathbf{Tf}\} = \mathbf{T}E\{\mathbf{f}\} = \mathbf{T}\eta_f, \tag{8.2.4}$$

а его корреляционная матрица

$$\mathbf{R}_p = E\{\mathbf{pp}^{*T}\} = E\{\mathbf{Tff}^{*T}\mathbf{T}^{*T}\} = \mathbf{TR}_f\mathbf{T}^{*T}. \tag{8.2.5}$$

Ковариационная матрица вектора $\mathbf{p}$

$$\mathbf{K}_p = \mathbf{TK}_f\mathbf{T}^{*T}. \tag{8.2.6}$$

Применение этих теоретических положений к операторам суперпозиции и унитарного преобразования описано в последующих главах.

8.3. ОПЕРАТОРЫ ПСЕВДООБРАЩЕНИЯ

При линейной обработке сигналов часто встречается задача «обращения» преобразования вида

$$\mathbf{p} = \mathbf{T}\mathbf{f} \tag{8.3.1}$$

с тем, чтобы выразить точное значение входного вектора $\mathbf{f}$ размера $Q \times 1$ или некоторую его оценку $\hat{\mathbf{f}}$ через выходной вектор $\mathbf{p}$ размера $P \times 1$. Если $\mathbf{T}$ — квадратная матрица, то очевидно, что

$$\hat{\mathbf{f}} = (\mathbf{T})^{-1}\mathbf{p}, \tag{8.3.2}$$

если обратная матрица существует. Если матрица $\mathbf{T}$ не квадратная, то для отыскания решения можно воспользоваться псевдообратной матрицей $\mathbf{T}^{+}$ размера $Q \times P$. В этом случае

$$\hat{\mathbf{f}} = \mathbf{T}^{+}\mathbf{p}. \tag{8.3.3}$$

Если решение существует и единственно, то «правильным» оператором псевдообращения будет тот, который обеспечивает точную оценку, т. е. $\hat{\mathbf{f}} = \mathbf{f}$. Это означает, что вектор $\mathbf{f}$ можно определить по наблюдаемому вектору $\mathbf{p}$ без ошибок. Если решение существует, но оно не единственно, то с помощью псевдообратного оператора можно выбрать решение с минимальной нормой. Если, наконец, точных решений не существует, то с помощью оператора псевдообращения можно найти наилучшее приближенное решение. Более подробно этот вопрос рассмотрен в последующих разделах. Разъяснения и доказательства многих из нижеприведенных положений содержатся в монографиях [4—6].

Первым из операторов псевдообращения будет рассмотрен оператор с обобщенной обратной матрицей $\mathbf{T}^{-}$, для которой выполняются следующие соотношения:

$$\mathbf{T}\mathbf{T}^{-} = (\mathbf{T}\mathbf{T}^{-})^{T}, \tag{8.3.4а}$$

$$\mathbf{T}^{-}\mathbf{T} = (\mathbf{T}^{-}\mathbf{T})^{T}, \tag{8.3.4б}$$

$$\mathbf{T}\mathbf{T}^{-}\mathbf{T} = \mathbf{T}, \tag{8.3.4в}$$

$$\mathbf{T}^{-}\mathbf{T}\mathbf{T}^{-} = \mathbf{T}^{-}. \tag{8.3.4г}$$

Обобщенная обратная матрица является единственной и при некоторых условиях ее можно записать в явном виде. Если $P > Q$, то систему уравнений (8.3.1) называют переопределенной, т. е. число компонент наблюдаемого вектора $\mathbf{p}$ превышает число подлежащих оценке компонент вектора $\mathbf{f}$. Если при этом ранг матрицы $\mathbf{T}$ равен Q, то обобщенная обратная матрица

$$\mathbf{T}^{-} = (\mathbf{T}^{T}\mathbf{T})^{-1}\mathbf{T}^{T}. \tag{8.3.5}$$

В противоположном случае, когда $P < Q$, систему (8.3.1) называют недоопределенной. Если при этом ранг матрицы $\mathbf{T}$ равен P, то обобщенная обратная матрица имеет вид

$$\mathbf{T}^{-} = \mathbf{T}^{T}(\mathbf{T}\mathbf{T}^{T})^{-1}. \tag{8.3.6}$$

Нетрудно показать, что матрицы, определенные соотношениями (8.3.5) и (8.3.6), удовлетворяют условиям (8.3.4). Если матрица $\mathbf{T}$ может быть представлена в виде прямого произведения (8.1.8), то обобщенная обратная матрица имеет вид

$$\mathbf{T}^{-} = \mathbf{T}_{C}^{-} \otimes \mathbf{T}_{R}^{-}, \tag{8.3.7}$$

где $\mathbf{T}_{R}^{-}$ и $\mathbf{T}_{C}^{-}$ — обобщенные обратные матрицы для линейных операторов обработки строк и столбцов. В этом случае сокращается объем вычислений, необходимых для обращения.

Еще один тип оператора псевдообращения имеет матрицу $\mathbf{T}^{S}$, называемую матрицей обращения методом наименьших квадратов, которая определяется следующими соотношениями:

$$\mathbf{T}\mathbf{T}^{S}\mathbf{T} = \mathbf{T}, \tag{8.3.8а}$$

$$\mathbf{T}\mathbf{T}^{S} = (\mathbf{T}\mathbf{T}^{S})^{T}. \tag{8.3.8б}$$

И наконец, условно обратная матрица $\mathbf{T}^{\#}$ определяется формулой

$$\mathbf{T}\mathbf{T}^{\#}\mathbf{T} = \mathbf{T}. \tag{8.3.9}$$

Анализ определений всех трех видов оператора псевдообращения показывает, что обобщенный обратный оператор является оператором обращения методом наименьших квадратов, а последний — оператором условного обращения. Для любого линейного оператора с матрицей $\mathbf{T}$ всегда существуют матрица обращения методом наименьших квадратов и условно обратная матрица, но они могут быть не единственными. Кроме того, для этих матриц обычно не удается найти явное выражение в конечной форме.

Ниже приведены некоторые полезные соотношения для матрицы $\mathbf{T}^{-}$, которая является обобщенной обратной матрицей относительно матрицы $\mathbf{T}$ размера $P \times Q$.

Обобщенное обращение транспонированной матрицы

$$[\mathbf{T}^{T}]^{-} = [\mathbf{T}^{-}]^{T}. \tag{8.3.10}$$

Обобщенное обращение обобщенной обратной матрицы

$$[\mathbf{T}^{-}]^{-} = \mathbf{T}. \tag{8.3.11}$$

Сохранение ранга

$$\text{Ранг } [\mathbf{T}^{-}] = \text{Ранг } [\mathbf{T}]. \tag{8.3.12}$$

Обобщенное обращение произведения матриц

$$[\mathbf{T}^{T}\mathbf{T}]^{-} = (\mathbf{T})^{-}(\mathbf{T}^{T})^{-}, \tag{8.3.13}$$

$$[\mathbf{A}\mathbf{B}]^{-} = \mathbf{B}^{-}\mathbf{A}^{-}, \tag{8.3.14}$$

где $\mathbf{A}$ и $\mathbf{B}$ — матрицы ранга R, имеющие размеры $P \times R$ и $R \times P$ соответственно.

Обобщенное обращение произведения ортогональных матриц

$$[\mathbf{ATB}]^{-} = \mathbf{B}^{T}\mathbf{T}^{-}\mathbf{A}^{T}, \tag{8.3.15}$$

где $\mathbf{A}$ и $\mathbf{B}$ — ортогональные матрицы размера $P \times P$ и $Q \times Q$ соответственно.

8.4. РЕШЕНИЕ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Систему линейных уравнений

$$\mathbf{p} = \mathbf{Tf}, \tag{8.4.1}$$

где $\mathbf{T}$ — матрица размера $P \times Q$, можно рассматривать как систему из P уравнений с Q неизвестными. Возможны три случая:

1. Система уравнений имеет единственное решение $\widehat{\mathbf{f}}$, для которого $\mathbf{T}\widehat{\mathbf{f}} = \mathbf{p}$.
2. Системе уравнений удовлетворяют несколько решений.
3. Система уравнений не имеет точного решения.

Если система уравнений имеет хотя бы одно решение, то ее называют совместной, а в противном случае — несовместной. Системы уравнений, не имеющие решений, часто получаются при исследовании физических систем, когда вектор $\mathbf{p}$ описывает последовательность измерений наблюдаемых величин, которые по предположению являются следствием воздействия некоторой недоступной для прямого наблюдения движущей силы, представленной вектором $\mathbf{f}$. Матрица $\mathbf{T}$ получается посредством математического моделирования характеристик реальной системы, для которой вектор $\mathbf{p}$ является выходным сигналом. При исправлении (реставрации) изображений вектор $\mathbf{f}$ обычно представляет исходное изображение, вектор $\mathbf{p}$ — смазанное изображение, а матрица $\mathbf{T}$ описывает дискретную математическую модель процесса, приводящего к смазыванию изображения. Поскольку матрица $\mathbf{T}$ и вектор $\mathbf{p}$ обычно определяются с некоторой погрешностью, может оказаться, что этот вектор не соответствует ни одному из возможных векторов воздействия $\mathbf{f}$.

Рассмотрим теперь вопрос о существовании решений системы уравнений $\mathbf{p} = \mathbf{Tf}$. Из структуры системы уравнений видно, что решение будет существовать тогда и только тогда, когда вектор $\mathbf{p}$ может быть получен линейной комбинацией столбцов матрицы $\mathbf{T}$. В этом случае говорят, что вектор $\mathbf{p}$ лежит в пространстве столбцов матрицы $\mathbf{T}$. Более строгой формулировкой условия существования решения является следующая [4]: система уравнений $\mathbf{p} = \mathbf{Tf}$ имеет решение тогда и только тогда, когда для матрицы $\mathbf{T}$ существует условно обратная матрица $\mathbf{T}^{\#}$, удовлетворяющая уравнению $\mathbf{TT}^{\#}\mathbf{p} = \mathbf{p}$.

Значит, для существования решения необходимо, чтобы при отображении из пространства наблюдаемых изображений в пространство исходных изображений (с помощью условно обратной матрицы $\mathbf{T}^{\#}$) и обратно (с помощью матрицы $\mathbf{T}$) снова получался вектор наблюдаемых величин $\mathbf{p}$. Если система уравнений является недоопределенной ($P < Q$), то решение существует тогда, когда ранг матрицы $\mathbf{T}$ равен P, т. е. полному числу строк. Во всех остальных случаях, в том числе и для переопределенной системы уравнений, существование решения необходимо проверять.

8.5. РЕШЕНИЕ СОВМЕСТНЫХ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Установив существование решения системы уравнений

$$\mathbf{p} = \mathbf{T}\mathbf{f}, \tag{8.5.1}$$

следует определить характер решения: является ли оно единственным или же решений несколько, а также какой вид имеет решение? Ответ на последний вопрос содержится в следующей фундаментальной теореме [4]:

Если решение системы уравнений $\mathbf{p} = \mathbf{T}\mathbf{f}$ существует, то в общем случае оно имеет вид

$$\hat{\mathbf{f}} = \mathbf{T}^{\#}\mathbf{p} + (\mathbf{I} - \mathbf{T}^{\#}\mathbf{T})\,\mathbf{v}, \tag{8.5.2}$$

где $\mathbf{T}^{\#}$ — матрица, условно обратная относительно матрицы $\mathbf{T}$, а $\mathbf{v}$ — произвольный вектор размера $Q \times 1$.

Для доказательства умножим обе части соотношения (8.5.2) на матрицу $\mathbf{T}$:

$$\mathbf{T}\hat{\mathbf{f}} = \mathbf{T}\mathbf{T}^{\#}\mathbf{p} + \mathbf{T}\,(\mathbf{I} - \mathbf{T}^{\#}\mathbf{T})\,\mathbf{v}. \tag{8.5.3}$$

Однако по условию существования решения $\mathbf{T}\mathbf{T}^{\#}\mathbf{p} = \mathbf{p}$. Кроме того, согласно определению условно обратной матрицы, $\mathbf{T}\,(\mathbf{I} - \mathbf{T}^{\#}\mathbf{T}) = \mathbf{0}$. Следовательно, $\widehat{\mathbf{T}\mathbf{f}} = \mathbf{p}$ и вектор $\hat{\mathbf{f}}$ является решением. Поскольку $\widehat{\mathbf{T}\mathbf{f}} = \mathbf{p}$, то, умножив обе части этого равенства на матрицу $\mathbf{T}^{\#}$, получим

$$\mathbf{T}^{\#}\mathbf{T}\hat{\mathbf{f}} = \mathbf{T}^{\#}\mathbf{p}, \tag{8.5.4а}$$

или

$$\mathbf{T}^{\#}\mathbf{p} - \mathbf{T}^{\#}\mathbf{T}\hat{\mathbf{f}} = \mathbf{0}. \tag{8.5.4б}$$

Прибавив к обеим частям вектор $\hat{\mathbf{f}}$, получим

$$\hat{\mathbf{f}} = \mathbf{T}^{\#}\mathbf{p} - \mathbf{T}^{\#}\mathbf{T}\hat{\mathbf{f}} + \hat{\mathbf{f}} = \mathbf{T}^{\#}\mathbf{p} + (\mathbf{I} - \mathbf{T}^{\#}\mathbf{T})\hat{\mathbf{f}}. \tag{8.5.5}$$

Этот результат совпадает с соотношением (8.5.2), если вектор $\hat{\mathbf{f}}$, стоящий в правой части формулы (8.5.5), заменить на произвольный вектор $\mathbf{v}$.

Поскольку обобщенная обратная матрица $\mathbf{T}^-$ и матрица $\mathbf{T}^S$ обращения методом наименьших квадратов являются условно обратными, то общее решение системы (8.5.1) также можно представить в виде

$$\hat{\mathbf{f}} = \mathbf{T}^S\mathbf{p} + (\mathbf{I} - \mathbf{T}^S\mathbf{T})\,\mathbf{v}, \tag{8.5.6а}$$

$$\hat{\mathbf{f}} = \mathbf{T}^-\mathbf{p} + (\mathbf{I} - \mathbf{T}^-\mathbf{T})\,\mathbf{v}. \tag{8.5.6б}$$

Решение очевидно, будет единственным, если $\mathbf{T}^{\#}\mathbf{T} = \mathbf{I}$. Во всех подобных случаях $\mathbf{T}^-\mathbf{T} = \mathbf{I}$. Исследсвав ранг матрицы $\mathbf{T}^-\mathbf{T}$, можно доказать, что [4] если решение системы уравнений $\mathbf{p} = \mathbf{T}\mathbf{f}$ существует, что оно единственно тогда и только тогда, когда ранг матрицы $\mathbf{T}$ размера $P \times Q$ равен Q.

Отсюда следует, что если решение недоопределенной системы уравнений существует, то оно не единственно. С другой стороны, переопределенная система уравнений может иметь только одно решение.

Пусть для системы уравнений (8.5.1) может быть получено точное решение. Рассмотрим оценку

$$\hat{\mathbf{f}} = \mathbf{T}^+\mathbf{p} = \mathbf{T}^+\mathbf{T}\mathbf{f}, \tag{8.5.7}$$

где $\mathbf{T}^+$ обозначает одну из матриц, псевдообратных относительно $\mathbf{T}$, которая не обязательно будет совпадать с этим решением, поскольку произведение матриц $\mathbf{T}^+\mathbf{T}$ может не равняться единичной матрице. Величину ошибки, т. е. отклонение оценки $\hat{\mathbf{f}}$ от истинного значения $\mathbf{f}$, обычно выражают через квадрат разности векторов $\mathbf{f}$ и $\hat{\mathbf{f}}$ в виде произведения

$$\mathscr{E}_E = (\mathbf{f} - \hat{\mathbf{f}})^T\,(\mathbf{f} - \hat{\mathbf{f}}), \tag{8.5.8а}$$

или как

$$\mathscr{E}_E = \operatorname{tr}\{(\mathbf{f} - \hat{\mathbf{f}})\,(\mathbf{f} - \hat{\mathbf{f}})^T\}. \tag{8.5.8б}$$

Подставив выражение (8.5.7) в (8.5.8а), получим

$$\mathscr{E}_E = \mathbf{f}^T\,[\mathbf{I} - (\mathbf{T}^+\mathbf{T})^T]\,[\mathbf{I} - (\mathbf{T}^+\mathbf{T})]\,\mathbf{f}. \tag{8.5.9}$$

Значение матрицы $\mathbf{T}^+$, при котором ошибка (8.5.8) оказывается минимальной, можно найти, приравняв нулю производную от ошибки $\mathscr{E}_E$ по вектору $\mathbf{f}$. Согласно соотношению (5.1.34),

$$\partial\mathscr{E}_E/\partial\mathbf{f} = \mathbf{0} = 2\,[\mathbf{I} - (\mathbf{T}^+\mathbf{T})^T]\,[\mathbf{I} - (\mathbf{T}^+\mathbf{T})]\,\mathbf{f}. \tag{8.5.10}$$

Равенство (8.5.10) удовлетворяется, если матрица $\mathbf{T}^+ = \mathbf{T}^-$, т. е. является обобщенной обратной матрицей относительно $\mathbf{T}$. При этом ошибка оценивания уменьшается до минимума, равного

$$\mathscr{E}_E = \mathbf{f}^T\,[\mathbf{I} - (\mathbf{T}^-\mathbf{T})]\,\mathbf{f}, \tag{8.5.11а}$$

или

$$\mathscr{E}_E = \mathbf{tr}\,\{\mathbf{f}\mathbf{f}^T[\mathbf{I} - (\mathbf{T}^-\mathbf{T})]\}. \tag{8.5.11б}$$

Как и ожидалось, ошибка становится равной нулю, когда $\mathbf{T}^-\mathbf{T} = \mathbf{I}$. Это произойдет, например, если обобщенная обратная матрица $\mathbf{T}^-$ имеет ранг Q и определяется соотношением (8.3.5).

8.6. ПРИБЛИЖЕННЫЕ РЕШЕНИЯ НЕСОВМЕСТНЫХ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Несовместность системы уравнений $\mathbf{p} = \mathbf{T}\mathbf{f}$ означает, что ни для одной из возможных оценок $\hat{\mathbf{f}}$ система не будет переходить в тождество при подстановке $\hat{\mathbf{f}}$ вместо $\mathbf{f}$. В таких случаях систему уравнений можно преобразовать к виду

$$\mathbf{p} = \mathbf{T}\mathbf{f} + \mathbf{e}\,(\mathbf{f}), \tag{8.6.1}$$

где $\mathbf{e}\,(\mathbf{f})$ — вектор ошибки, зависящий от $\mathbf{f}$. Найдем теперь такое значение оценки $\hat{\mathbf{f}}$, при котором оказывается минимальной величина ошибки, выражаемая двумя эквивалентными соотношениями:

$$\mathscr{E}_M = [\mathbf{e}\,(\hat{\mathbf{f}})]^T\,[\mathbf{e}\,(\hat{\mathbf{f}})] = [\mathbf{p} - \mathbf{T}\hat{\mathbf{f}}]^T\,[\mathbf{p} - \mathbf{T}\hat{\mathbf{f}}], \tag{8.6.2а}$$

$$\mathscr{E}_M = \mathbf{tr}\,\{[\mathbf{e}\,(\hat{\mathbf{f}})]\,[\mathbf{e}\,(\hat{\mathbf{f}})]^T\} = \mathbf{tr}\,\{[\mathbf{p} - \mathbf{T}\hat{\mathbf{f}}]\,[\mathbf{p} - \mathbf{T}\hat{\mathbf{f}}]^T\}. \tag{8.6.2б}$$

Пусть символ $\mathbf{T}^+$ обозначает псевдообратную матрицу, с помощью которой получается оценка

$$\hat{\mathbf{f}} = \mathbf{T}^+\mathbf{p}. \tag{8.6.3}$$

Прибавив и отняв произведение $\mathbf{T}\mathbf{T}^+\mathbf{p}$ внутри обеих скобок соотношения (8.6.2а), получим

$$\mathscr{E}_M = [(\mathbf{I} - \mathbf{T}\mathbf{T}^+)\,\mathbf{p} + \mathbf{T}\,(\mathbf{T}^+\mathbf{p} - \hat{\mathbf{f}})]^T\,[(\mathbf{I} - \mathbf{T}\mathbf{T}^+)\,\mathbf{p} + \mathbf{T}\,(\mathbf{T}^+\mathbf{p} - \hat{\mathbf{f}})]. \tag{8.6.4}$$

После перемножения имеем

$$\mathscr{E}_M = [(\mathbf{I} - \mathbf{T}\mathbf{T}^+)\,\mathbf{p}]^T\,[(\mathbf{I} - \mathbf{T}\mathbf{T}^+)\,\mathbf{p}] + [\mathbf{T}\,(\mathbf{T}^+\mathbf{p} - \hat{\mathbf{f}})]^T\,[\mathbf{T}\,(\mathbf{T}^+\mathbf{p} - \hat{\mathbf{f}})] + [(\mathbf{I} - \mathbf{T}\mathbf{T}^+)\,\mathbf{p}]^T\,[\mathbf{T}\,(\mathbf{T}^+\mathbf{p} - \hat{\mathbf{f}})] + [\mathbf{T}\,(\mathbf{T}^+\mathbf{p} - \hat{\mathbf{f}})]^T\,[(\mathbf{I} - \mathbf{T}\mathbf{T}^+)\,\mathbf{p}]. \tag{8.6.5}$$

Два перекрестных члена будут равны нулю, если $\mathbf{T}\mathbf{T}^+\mathbf{T} = \mathbf{T}$ и $(\mathbf{T}\mathbf{T}^+) = (\mathbf{T}\mathbf{T}^+)^T$. Однако при выполнении этих условий матрица $\mathbf{T}^+$ является матрицей обращения методом наименьших квадратов, т. е. $\mathbf{T}^+ = \mathbf{T}^S$. Тогда ошибка будет равна сумме двух положительных слагаемых:

$$\mathscr{E}_M = [(\mathbf{I} - \mathbf{T}\mathbf{T}^S)\,\mathbf{p}]^T\,[(\mathbf{I} - \mathbf{T}\mathbf{T}^S)\,\mathbf{p}] + [\mathbf{T}\,(\mathbf{T}^S\mathbf{p} - \hat{\mathbf{f}})]^T\,[\mathbf{T}\,(\mathbf{T}^S\mathbf{p} - \hat{\mathbf{f}})]. \tag{8.6.6}$$

Второй член равенства (8.6.6) превращается в нуль, так как $\hat{\mathbf{f}} = \mathbf{T}^S\mathbf{p}$. Следовательно, ошибка уменьшается до величины

$$\mathscr{E}_M = \mathbf{p}^T[\mathbf{I} - \mathbf{T}\mathbf{T}^S]\,\mathbf{p}, \tag{8.6.7а}$$

или, что то же самое,

$$\mathscr{E}_M = \mathrm{tr}\,[\mathbf{p}\mathbf{p}^T(\mathbf{I} - \mathbf{T}\mathbf{T}^S)]. \tag{8.6.7б}$$

Как и ожидалось, ошибка равна нулю, если $\mathbf{T}\mathbf{T}^S = \mathbf{I}$.

Решение, полученное псевдообращением по методу наименьших квадратов, может быть не единственным. Если при псевдообращении ввести дополнительные условия $\mathbf{T}^+\mathbf{T}\mathbf{T}^+ = \mathbf{T}$ и $\mathbf{T}^+\mathbf{T} = (\mathbf{T}^+\mathbf{T})^T$, при которых матрица $\mathbf{T}^+$ является обобщенной обратной (т. е. $\mathbf{T}^+ = \mathbf{T}^-$), то можно показать, что оценка, полученная с помощью этой матрицы ($\hat{\mathbf{f}} = \mathbf{T}^-\mathbf{g}$), является решением с минимальной нормой в том смысле, что

$$\hat{\mathbf{f}}^T\hat{\mathbf{f}} \leqslant \tilde{\mathbf{f}}^T\tilde{\mathbf{f}}, \tag{8.6.8}$$

где $\tilde{\mathbf{f}}$ — произвольная оценка, найденная методом наименьших квадратов. Если обобщенная обратная матрица $\mathbf{T}^-$ имеет ранг Q и удовлетворяет определению (8.3.5), то произведение $\mathbf{T}\mathbf{T}^-$ не обязательно равно единичной матрице, а ошибку можно найти из соотношений (8.6.7). Если же матрица $\mathbf{T}^-$ имеет ранг P, т. е. соответствует определению (8.3.6), то $\mathbf{T}\mathbf{T}^- = \mathbf{I}$ и ошибка равна нулю.

В последующих главах будет показано, как данные теоретические положения применяются для исправления, анализа и кодирования изображений.

ЛИТЕРАТУРА

1. Pratt W. K , Vector Formulation of Two Dimensional Signal Processing Operations, *J. Computer Graphics and Image Processing*, **4**, 1, 1—24 (March 1975), Academic Press, New York.
2. Eklundh J. O., A Fast Computer Method for Matrix Transposing, *IEEE Trans. Computers*, **C-21**, 7, 801—803 (July 1972).
3. Twogood R. E., Ekstrom M. P., An Extension of Eklundh's Matrix Transposition Algorithm and Its Applications in Digital Image Processing, *IEEE Trans. Computers*, **C-25**, 9, 950—952 (September 1976).
4. Graybill F. A., Introduction to Matrices with Applications in Statistics, Wadsworth, Belmont, Cal , 1969.
5. Rao C. R., Mitra S. K., Generalized Inverse of Matrices and Its Applications, Wiley, New York, 1971.
6. Albert A., Regression and the Moore — Penroce Pseudoinverse, Academic Press, New York, 1972. [Имеется перевод: Алберт А., Регрессия, псевдоинверсия и рекуррентное оценивание. — М.: Наука, 1977.]

Глава 9

ОПЕРАТОР СУПЕРПОЗИЦИИ

Двумерный дискретный оператор суперпозиции определяется, во-первых, как оператор дискретной линейной фильтрации, применяемый к конечному массиву отсчетов, и, во-вторых, как дискретизованный оператор аналоговой фильтрации. Обычно эти два определения не эквивалентны. Ниже будут применяться оба определения. Кроме того, будет введен дискретный оператор суперпозиции с циклически (периодически) продолженной областью действия и описаны соотношения между тремя операторами.

9.1. ОПЕРАТОР СУПЕРПОЗИЦИИ КОНЕЧНЫХ МАССИВОВ

Рассмотрим сначала дискретный оператор суперпозиции конечного массива отсчетов [1] $F(n_1, n_2)$ (где $n_1, n_2 = 1, 2, \ldots, N$) с конечным массивом $H(l_1, l_2; m_1, m_2)$ (где $l_1, l_2 = 1, 2, \ldots, L$), играющим роль импульсного отклика. В общем случае импульсный отклик может изменяться в зависимости от координат (m_1, m_2) отсчета в выходном массиве Q. Операция суперпозиции в ограниченной области определяется соотношением

$$Q(m_1, m_2) = \sum_{n_1=1}^{m_1} \sum_{n_2=1}^{m_2} F(n_1, n_2) H(m_1 - n_1 + 1, m_2 - n_2 + 1; m_1, m_2), \qquad (9.1.1)$$

где $m_1, m_2 = 1, 2, \ldots, M$, а массивы H и F имеют нулевые значения вне областей изменения соответствующих индексов. Анализируя предельные значения индексов отсчетов импульсного отклика, можно убедиться, что $M = N + L - 1$ и, следовательно, выходной массив Q имеет большие размеры, чем исходный (рис. 9.1.1).

Если массивы F и Q представлены соответственно в виде вектора $\mathbf{f}$ размера $N^2 \times 1$ и вектора $\mathbf{q}$ размера $M^2 \times 1$, то преобразование (9.1.1) можно записать как [1]

$$\mathbf{q} = \mathbf{D}\mathbf{f}, \qquad (9.1.2)$$

---

[1]) Для простоты обозначений принято, что все массивы отсчетов квадратные.

где $\mathbf{D}$ — матрица размера $M^2 \times N^2$, содержащая отсчеты импульсного отклика. Матрицу $\mathbf{D}$ оператора суперпозиции удобно раз-

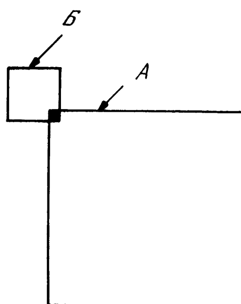


Рис. 9.1.1. Суперпозиция конечных массивов отсчетов импульсного отклика и исходного изображения.

А — массив из $N \times N$ отсчетов изображения; *Б* — повернутый на 180° массив из $L \times L$ отсчетов импульсного отклика

делить на блоки $\mathbf{D}_{m,n}$ размера $M \times N$. Проанализировав пределы суммирования в выражении (9.1.1), можно показать, что

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{1,1} & 0 & & & 0 \\ \mathbf{D}_{2,1} & \mathbf{D}_{2,2} & & & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ \mathbf{D}_{L,1} & \mathbf{D}_{L,2} & & & \mathbf{D}_{M-L+1,N} \\ 0 & \mathbf{D}_{L+1,2} & & & \vdots \\ 0 & 0 & & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{D}_{M,N} \end{bmatrix}. \qquad (9.1.3)$$

Произвольный ненулевой элемент матрицы $\mathbf{D}$ имеет вид

$$D_{m_2,\,n_2}(m_1,\ n_1) = H(m_1 - n_1 + 1,\ m_2 - n_2 + 1;\ m_1,\ m_2), \qquad (9.1.4)$$

где $1 \leqslant n_i \leqslant N$, а $n_i \leqslant m_i \leqslant n_i + L - 1$. Отсюда следует, что матрица $\mathbf{D}$ имеет регулярную структуру и заполнена довольно редко, причем ненулевые блоки, группирующиеся в виде полосы в средней части матрицы $\mathbf{D}$, содержат зоны нулевых элементов.

Если форма импульсного отклика инвариантна относительно сдвига (т. е. одинакова для всех точек выходного массива), то структура матрицы $\mathbf{D}$ не зависит в явной форме от координат $(m_1,\ m_2)$ выходного отсчета. Тогда

$$\mathbf{D}_{m_2,\,n_2} = \mathbf{D}_{m_2+1,\ n_2+1}. \qquad (9.1.5)$$

Таким образом, все столбцы матрицы **D** образуются сдвигом первого столбца. В этом случае оператор суперпозиции называется оператором свертки конечных массивов. На рис. 9.1.2,*а* приведены полученные на ЦВМ распечатки матриц, фигурирующих в операции свертки конечных массивов, для случая, когда входной массив имеет размеры 2×2 ($N = 2$), выходной массив 4×4 ($M = 4$), а массив отсчетов импульсного отклика 3×3 ($L = 3$). Пары целых

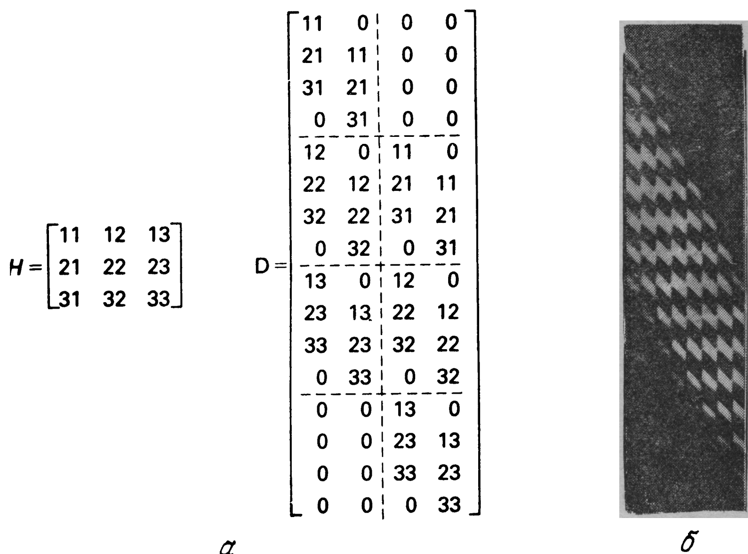


Рис. 9.1.2. Примеры матриц операторов свертки конечных массивов.
а — общий случай, $M = 4$, $N = 2$, $L = 3$, *б* — импульсный отклик гауссовой формы, $M = 16$, $N = 8$, $L = 9$

чисел (i, j) [1] в матрице **D** обозначают (i, j)-й элемент матрицы **H** (i, j). Структура матрицы **D** лучше видна на примере матрицы большего размера, показанной на рис. 9.1.2, *б*. В рассматриваемой матрице $M = 16$, $N = 8$, $L = 9$, а импульсный отклик симметричен и имеет гауссову форму. Заметим, что в данном примере размеры матрицы **D** равны 256×64.

Пользуясь методикой, примененной при выводе соотношения (8.1.7), оператор суперпозиции можно представить в матричной форме

$$\mathbf{Q} = \sum_{m=1}^{M} \sum_{n=1}^{N} \mathbf{D}_{m,\,n}\mathbf{F}\mathbf{v}_n\mathbf{u}_m^T. \qquad (9.1.6)$$

---

[1] В этих парах отброшены запятые и пара индексов записывается как двузначное число. — *Прим. перев.*

Если импульсный отклик является инвариантным относительно сдвига и разделимым, т. е.

$$\mathbf{H} = \mathbf{h}_C \mathbf{h}_R^T, \tag{9.1.7}$$

где $\mathbf{h}_C$ и $\mathbf{h}_R$ — вектор-столбцы, описывающие соответственно характер изменения импульсного отклика по столбцам и строкам, то

$$\mathbf{D} = \mathbf{D}_C \otimes \mathbf{D}_R, \tag{9.1.8}$$

Матрицы $\mathbf{D}_R$ и $\mathbf{D}_C$ имеют размеры $M \times N$ и структуру вида

$$\mathbf{D}_R = \begin{bmatrix} h_R(1) & 0 & \cdots & 0 \\ h_R(2) & h_R(1) & & \vdots \\ h_R(3) & h_R(2) & \ddots & 0 \\ & & & h_R(1) \\ \vdots & & & \vdots \\ h_R(L) & & & \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & \cdots & 0 & h_R(L) \end{bmatrix}. \qquad (9.1.9)$$

Операция двумерной свертки в этом случае сводится к последовательному вычислению одномерных сверток по строкам и столбцам. Таким образом,

$$\mathbf{Q} = \mathbf{D}_C \mathbf{F} \mathbf{D}_R^T. \tag{9.1.10}$$

Для получения конечной свертки или суперпозиции в общем случае необходимо выполнить N^2L^2 арифметических операций, причем в это число не входят умножения на нулевые элементы матрицы $\mathbf{D}$. Если же оператор является разделимым, т. е. удовлетворяет равенству (9.1.10), то достаточно выполнить $NL\,(M + N)$ операций.

9.2. ДИСКРЕТИЗОВАННЫЙ ОПЕРАТОР СУПЕРПОЗИЦИИ

При цифровой обработке изображений во многих случаях требуется перевести в дискретную форму интегральный оператор суперпозиции, связывающий непрерывные изображения на входе и выходе линейной системы. Такие явления, как смазывание изображения, вызванное несовершенством оптической системы, апертурные искажения или искажения при наблюдении через турбулентную атмосферу, можно описать с помощью следующего интегрального уравнения:

$$\tilde{G}(x, y) = \int\limits_{-\infty}^{\infty}\!\!\int \tilde{F}(\alpha, \beta)\, \tilde{J}(x, y; \alpha, \beta)\, d\alpha\, d\beta, \tag{9.2.1a}$$

где $\tilde{F}(x, y)$ и $\tilde{G}(x, y)$ описывают соответственно входное и выходное изображения, а ядро $\tilde{J}(x, y;\ \alpha, \beta)$ представляет импульсный отклик линейной системы [1]). Импульсный отклик может быть функцией четырех переменных — координат точек входной и выходной плоскостей. Если линейная система является пространственно-инвариантной, то изображение на ее выходе можно описать с помощью интеграла свертки

$$\tilde{G}(x, y) = \iint\limits_{-\infty}^{\infty} \tilde{F}(\alpha, \beta)\, \tilde{J}(x-\alpha, y-\beta)\, d\alpha\, d\beta. \qquad (9.2.1б)$$

При дискретной обработке выходное изображение будет представлено отсчетами. Поэтому интеграл суперпозиции или свертки следует преобразовать так, чтобы явно выразить связь отсчетов выходного изображения со значениями функции, описывающей входное изображение. Вопрос о дискретном представлении интегральных преобразований важен, поскольку получающиеся при этом ошибки могут привести к большим ошибкам в обработанном изображении или даже к потере устойчивости процесса цифровой обработки. Кроме того, от выбора алгоритма дискретного представления, как правило, сильно зависит сложность самой обработки.

На первом этапе процесса дискретизации интеграла суперпозиции выходное изображение дискретизуется с помощью решетки из $(2J+1) \times (2J+1)$ дельта-импульсов Дирака, разделенных интервалом ΔS. В результате образуется массив отсчетов [2])

$$\tilde{G}(j_1 \Delta S,\ j_2 \Delta S) = \tilde{G}(x, y)\, \delta(x - j_1 \Delta S,\ y - j_2 \Delta S), \qquad (9.2.2)$$

где $-J \leqslant j_i \leqslant J$. Дельта-функцию можно ввести в подынтегральное выражение интеграла (9.2.1a); в результате получается соотношение

$$\tilde{G}(j_1 \Delta S,\ j_2 \Delta S) = \iint\limits_{-\infty}^{\infty} \tilde{F}(\alpha, \beta)\, \tilde{J}(j_1 \Delta S,\ j_2 \Delta S;\ \alpha, \beta)\, d\alpha\, d\beta. \qquad (9.2.3)$$

Следует отметить, что дискретизация выходного изображения осуществляется по координатам плоскости наблюдаемого изображения (x, y) и не затрагивает текущие переменные (α, β), по которым проводится интегрирование.

---

[1]) В пределах данной главы знак $\sim$ над символом функции указывает, что соответствующие пространственные аргументы принимают значения обоих знаков, т. е. в интервалы их изменения входят отрезки как положительной, так и отрицательной полуосей.

[2]) Для упрощения обозначений принято, что шаги дискретизации по горизонтали и вертикали одинаковы. Эффекты, связанные с конечными размерами дискретизирующих импульсов, учесть нетрудно. Для этого импульсный отклик следует заменить сверткой $\tilde{J}(x, y;\ \alpha, \beta) * P(-x, -y)$, где функция $P(x, y)$ описывает форму дискретизирующего импульса.

На следующем этапе импульсный отклик следует ограничить по длительности. Для этого положим

$$\tilde{J}(x, y; \alpha, \beta) = 0 \qquad (9.2.4)$$

при $|x - \alpha| > T$ и $|y - \beta| > T$. Тогда

$$\tilde{G}(j_1 \Delta S, j_2 \Delta S) = \int\limits_{j_1 \Delta S - T}^{j_1 \Delta S + T} \int\limits_{j_2 \Delta S - T}^{j_2 \Delta S + T} \tilde{F}(\alpha, \beta) \tilde{J}(j_1 \Delta S, j_2 \Delta S; \alpha, \beta) \, d\alpha \, d\beta. \qquad (9.2.5)$$

Следует отметить, что усечение импульсного отклика эквивалентно умножению его на весовую функцию («окно») $V(x, y)$, равную единице внутри квадрата $|x| < T$, $|y| < T$ и нулю за его пределами. Согласно теореме о спектре свертки, спектр получаемого после усечения изображения равен свертке спектров изображения $G(x, y)$ и весовой функции (окна) $V(x, y)$, причем спектр последней есть двумерная sinc-функция. Такое искажение спектра $G(x, y)$ приводит к появлению паразитных компонент с высокими пространственными частотами, группирующихся вблизи частот, кратных $2\pi/T$ (явление Гиббса). Искажения, вносимые при усечении, можно уменьшить, используя более совершенные весовые функции, такие, как окна Бартлетта, Блэкмана, Хемминга или Ханна[1] [3], которые позволяют сгладить нежелательные эффекты, связанные с наличием разрывов прямоугольного окна. Выбор подходящего окна имеет большое значение, поскольку некорректность интегрального преобразования (9.2.1а) может привести к резкому усилению искажений, вызванных усечением импульсного отклика в процессе восстановления изображения.

Следующий этап процесса дискретизации состоит в том, что непрерывное исходное изображение $\tilde{F}(\alpha, \beta)$ представляется набором значений описывающей его функции в узлах прямоугольной сетки с шагом ΔI и размерами $(2K + 1) \times (2K + 1)$. Эта операция является не взятием реальных отсчетов, а всего лишь математическим преобразованием, в результате которого получается массив

$$\tilde{F}(k_1 \Delta I, k_2 \Delta I) = \tilde{F}(\alpha, \beta) \, \delta(\alpha - k_1 \Delta I, \beta - k_2 \Delta I), \qquad (9.2.6)$$

где $K_{iL} \leqslant k_i \leqslant K_{iU}$, причем K_{iL} и K_{iU} обозначают минимальное и максимальное значения индекса k_i.

---

[1] Иногда весовую функцию Ханна. (Г. von Hann) ошибочно называют «окном Хеннинга». Хеннинг (hanning) — операция умножения на эту весовую функцию. — *Прим. перев.*

Если конечной целью является оценка непрерывного исходного изображения путем обработки реальных отсчетов наблюдаемого физического поля, то шаг ΔI должен быть взят достаточно малым, чтобы для исходного изображения удовлетворялся критерий Найквиста. Это значит, что если исходное изображение имеет спектр конечной ширины, то расстояние между узлами следует положить равным соответствующему найквистовскому интервалу. В идеальном случае это обеспечивает возможность точного восстановления изображения $\tilde{F}(x, y)$ посредством интерполяции оцененных значений $\tilde{F}(k_1 \Delta I, k_2 \Delta I)$.

Воспользовавшись формулами для приближенного вычисления интегралов, в равенстве (9.2.5) интеграл можно аппроксимировать суммой [2]. В этом случае реальные отсчеты изображения можно описать выражением

$$\tilde{G}(j_1 \Delta S, j_2 \Delta S) = \sum_{k_1=K_{1L}}^{K_{1U}} \sum_{k_2=K_{2L}}^{K_{2U}} \tilde{F}(k_1 \Delta I, k_2 \Delta I) \tilde{W}(k_1, k_2) \tilde{J}(j_1 \Delta S, j_2 \Delta S; k_1 \Delta I, k_2 \Delta I), \quad (9.2.7)$$

где $\tilde{W}(k_1, k_2)$ — весовые коэффициенты в конкретной приближенной формуле, примененной для вычислений. Обычно пользуются формулой прямоугольников, для которой все весовые коэффициенты равны единице. В любом случае удобно объединить весовые коэффициенты и отсчеты импульсного отклика, так что

$$\tilde{H}(j_1 \Delta S, j_2 \Delta S; k_1 \Delta I, k_2 \Delta I) \equiv \equiv \tilde{W}(k_1, k_2) \tilde{J}(j_1 \Delta S, j_2 \Delta S; k_1 \Delta I), k_2 \Delta I). \quad (9.2.8)$$

Тогда

$$\tilde{G}(j_1 \Delta S, j_2 \Delta S) = \sum_{k_1=K_{1L}}^{K_{1U}} \sum_{k_2=K_{2L}}^{K_{2U}} \tilde{F}(k_1 \Delta I, k_2 \Delta I) \tilde{H}(j_1 \Delta S, j_2 \Delta S; k_1 \Delta I, k_2 \Delta I). \quad (9.2.9)$$

Следует отметить, что функция $\tilde{H}$ не подвергалась дискретизации по двум первым переменным. В формулах просто фигурируют ее значения в соответствующих точках. Пределы суммирования в равенстве (9.2.9) равны

$$K_{iL} = \left[j_i \frac{\Delta S}{\Delta I} - \frac{T}{\Delta I} \right]_N, \quad K_{iU} = \left[j_i \frac{\Delta S}{\Delta I} + \frac{T}{\Delta I} \right]_N, \quad (9.2.10)$$

где символ $[\cdot]_N$ обозначает округление до ближайшего целого числа.

На рис. 9.2.1 приведен пример того, как реальные отсчеты $\tilde{G}(j_1 \Delta S, j_2 \Delta S)$ выходного изображения сопоставляются с узловыми

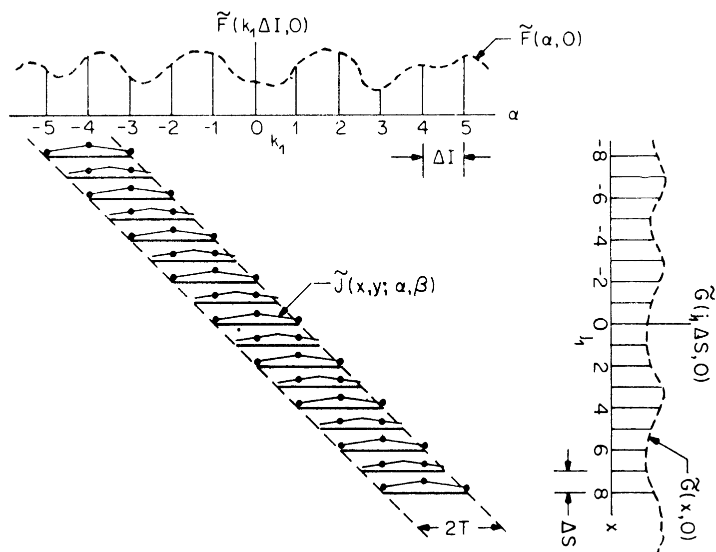


Рис. 9.2.1. Соотношение между отсчетами реального изображения и значениями исходного изображения в узловых точках при численном представлении интеграла суперпозиции.

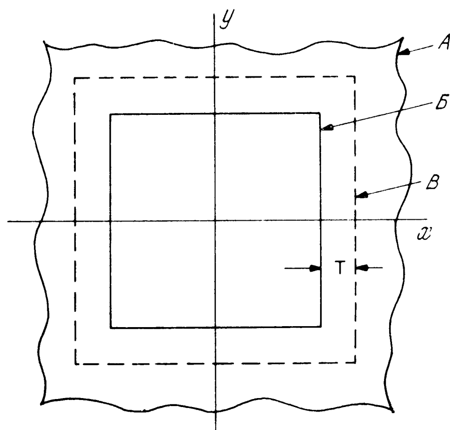


Рис. 9.2.2. Соотношение между областями расположения реальных отсчетов и узловых точек при численном представлении интеграла суперпозиции.

А — исходное изображение; *Б* — область расположения отсчетов наблюдаемого изображения; *В* — область расположения узловых точек

точками $F(k_1\,\Delta I,\ k_2\,\Delta I)$ исходного изображения. В данном примере интервал между узлами выбран вдвое больше расстояния между реальными отсчетами. Значения импульсного отклика, использованные при вычислении суммы (9.2.9), показаны на этом рисунке точками.

В связи с дискретным представлением (9.2.9) линейного интегрального преобразования следует сделать важное замечание: независимо от соотношения между числом узловых точек и числом реальных отсчетов площадь исходного изображения $\tilde{F}(x,\ y)$, на которой расположены узловые точки, влияющие на значения реальных отсчетов, всегда оказывается больше площади дискретизованного выходного изображения $\tilde{G}(j_1\,\Delta S,\ j_2\,\Delta S)$. Как показано на рис. 9.2.2, размеры обоих изображений с точностью до одного интервала дискретизации связаны соотношением

$$J\,\Delta S + T = K\,\Delta I. \tag{9.2.11}$$

Итак, построена дискретная конечная модель линейного интегрального преобразования, которая связывает отсчеты выходного изображения $\tilde{G}(j_1\,\Delta S,\ j_2\,\Delta S)$ со значениями элементов исходного изображения $\tilde{F}(k_1\,\Delta I,\ k_2\,\Delta I)$ посредством математической операции дискретного линейного преобразования. Такая операция является аппроксимацией непрерывного преобразования, поскольку импульсный отклик $\tilde{J}(x,\ y;\ \alpha,\ \beta)$ подвергался усечению, а интеграл вычислялся по приближенным квадратурным формулам. Ясно, что, расширяя область определения импульсного отклика, ошибку усечения можно сделать сколь угодно малой, хотя это и связано с увеличением объема вычислений. Ошибки за счет приближенного вычисления интегралов можно уменьшить, применяя более точные квадратурные формулы, но также ценой усложнения вычислений. Следует, однако, отметить, что дискретный вариант линейного интегрального преобразования является точной аппроксимацией последнего, если все функции пространственных переменных, входящие в формулу (9.2.1), имеют спектры ограниченной ширины, а интервалы при дискретизации выходного изображения и численном представлении исходного изображения выбраны в соответствии с критерием Найквиста [4]. Вопрос о точности и устойчивости дискретного представления рассмотрен в гл. 14.

Часто оказывается, что пользоваться соотношением (9.2.9) удобнее, если оно представлено в векторной форме. С этой целью нумерацию элементов массивов $\tilde{G}$ и $\tilde{F}$ следует изменить так, чтобы образовались соответственно массивы размерами $M\times M$ и $N\times N$ с положительными индексами. Допустим, что

$$F(n_1\,\Delta I,\ n_2\,\Delta I) = \tilde{F}(k_1\,\Delta I,\ k_2\,\Delta I), \tag{9.2.12a}$$

где $n_i = k_i + K + 1$, а

$$G(m_1 \Delta S,\ m_2 \Delta S) = \tilde{G}(j_1 \Delta S,\ j_2 \Delta S), \qquad (9.2.12б)$$

где $m_i = j_i + J + 1$. Кроме того, определим импульсный отклик так, что

$$H(m_1 \Delta S,\ m_2 \Delta S;\ n_1 \Delta I,\ n_2 \Delta I) =$$
$$= \tilde{H}(j_1 \Delta S,\ j_2 \Delta S;\ k_1 \Delta I,\ k_2 \Delta I). \qquad (9.2.12в)$$

На рис. 9.2.3 показаны геометрические соотношения между этими функциями.

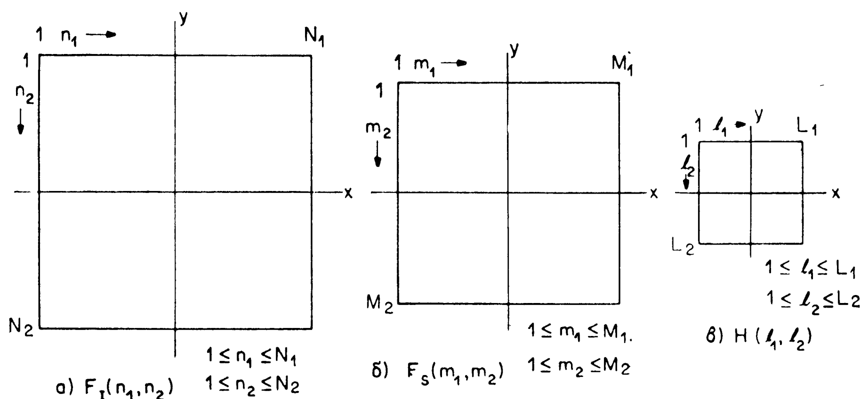


Рис. 9.2.3. Массивы отсчетов изображения и импульсного отклика.

Дискретное преобразование (9.2.9) применительно к сдвинутым массивам имеет следующий вид:

$$G(m_1 \Delta S,\ m_2 \Delta S) = \sum_{n_1=N_{1L}}^{N_{1U}} \sum_{n_2=N_{2L}}^{N_{2U}} F(n_1 \Delta I,\ n_2 \Delta I)\, H(m_1 \Delta S,$$
$$m_2 \Delta S;\ n_1 \Delta I,\ n_2 \Delta I), \qquad (9.2.13)$$

где $1 \leqslant m_i \leqslant M$, а

$$N_{iL} = \left[m_i \frac{\Delta S}{\Delta I}\right]_N, \qquad N_{iU} = \left[m_i \frac{\Delta S}{\Delta I} + \frac{2T}{\Delta I}\right]_N.$$

Воспользовавшись методикой, описанной в гл. 5, можно сформировать векторы **g** и **f**, являющиеся разверткой матриц **G** и **F** по столбцам. Эти векторы связаны соотношением [1]

$$\mathbf{g} = \mathbf{B}\mathbf{f}, \qquad (9.2.14)$$

где $\mathbf{B}$ — матрица размера $M^2 \times N^2$, имеющая следующую структуру:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,L} & 0 & \cdots & 0 \\ 0 & \ddots & & & \ddots & & \vdots \\ \vdots & & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{B}_{M,N-L+1} & & \cdots & \mathbf{B}_{M,N} \end{bmatrix}. \quad (9.2.15)$$

Элементы матрицы $\mathbf{B}$ определяются равенством

$$B_{m_2,\, n_2}(m_1,\ n_1) = H(m_1\,\Delta S,\ m_2\,\Delta S;\ n_1\,\Delta I,\ n_2\,\Delta I) \quad (9.2.16)$$

при $1 \leqslant m_i \leqslant M$, $m_i \leqslant n_i \leqslant m_i + L - 1$, где $L = [2T/\Delta I\,]_N + 1$ есть нечетное целое число, ближайшее к размерам импульсного отклика, выраженным числом шагов дискретизации ΔI. В дальнейшем для простоты матрицу $\mathbf{B}$ будем называть матрицей смазывания. Если импульсный отклик инвариантен относительно смещения, т. е.

$$H(j_1\,\Delta S,\ j_2\,\Delta S;\ k_1\,\Delta I,\ k_2\,\Delta I) = \\ = H(j_1\,\Delta S - k_1\,\Delta I j_2\,\Delta S - k_2\,\Delta I), \quad (9.2.17)$$

то операция дискретной суперпозиции (9.2.13) превращается в дискретную свертку:

$$G(m_1\,\Delta S,\ m_2\,\Delta S) = \sum_{n_1=N_{1L}}^{N_{1U}} \sum_{n_2=N_{2L}}^{N_{2U}} F(n_1\,\Delta I,\ n_2\,\Delta I)\, H(m_1\,\Delta S - \\ - n_1\,\Delta I + 2T,\ m_2\,\Delta S - n_2\,\Delta I + 2T). \quad (9.2.18)$$

Интересен частный случай, когда $\Delta S/\Delta I = k$, где k — целое число. При этом

$$G(m_1\,\Delta S,\ m_2\,\Delta S) = \sum_{n_1=m_1k}^{m_1k+L-1} \sum_{n_2=m_2k}^{m_2k+L-1} F(n_1\,\Delta I,\ n_2\,\Delta I)\, H[m_1k\,\Delta I - \\ - n_1\,\Delta I + (L-1)\,\Delta I,\ m_2k\,\Delta I - n_2\,\Delta I + (L-1)\,\Delta I]. \quad (9.2.19)$$

Теперь элементы матрицы смазывания равны

$$B_{m_2,\, n_2}(m_1,\ n_1) = H[m_1k\,\Delta I - n_1\,\Delta I + (L-1)\,\Delta I, \\ m_2k\,\Delta I - n_2\,\Delta I + (L-1)\,\Delta I], \quad (9.2.20)$$

где $1 \leqslant m_i \leqslant M$ и $m_i \leqslant n_i \leqslant m_i + L - 1$. Далее, если $\Delta S = \Delta I$, то элементы матрицы смазывания принимают вид

$$B_{m_2,\, n_2}(m_1,\ n_1) = H(m_1 - n_1 + L,\ m_2 - n_2 + L). \quad (9.2.21)$$

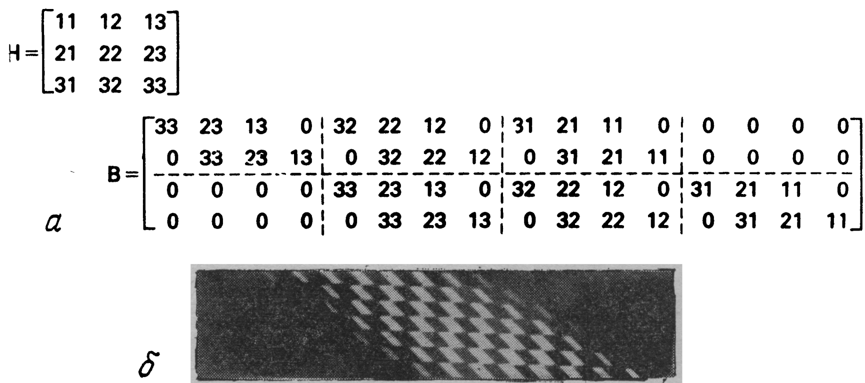
Кроме того,

$$\mathbf{B}_{m_2,\, n_2} = \mathbf{B}_{m_2+1,\, n_2+1}. \quad (9.2.22)$$

Следовательно, все строки матрицы **B** получаются сдвигом первой строки. Тогда оператор **B** является дискретизованным оператором свертки, а представление суммы (9.2.19) упрощается:

$$G(m_1,\ m_2) = \sum_{n_1=m_1}^{m_1+L-1} \sum_{n_2=m_2}^{m_2+L-1} F(n_1,\ n_2)\, H(m_1 - n_1 + L,\ m_2 - n_2 + L). \tag{9.2.23}$$

На рис. 9.2.4, *а* приведена распечатка матриц дискретизованного оператора свертки, когда входной массив состоит из 4×4 отсчетов ($N=4$), импульсный отклик представлен 3×3 отсчетами ($L=3$), а выходной массив содержит 2×2 отсчета ($M=2$). На рис. 9.2.4, *б* показан вид матрицы **B** для случая более высоких размерностей, когда $M = 8$, $N = 16$, $L = 9$, а импульсный отклик имеет гауссову форму.



$$\mathbf{H} = \begin{bmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{bmatrix}$$

$$\mathbf{B} = \left[\begin{array}{cccc|cccc|cccc|cccc} 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 & 0 \\ 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 \end{array}\right]$$

Рис. 9.2.4. Примеры матриц дискретизованных операторов свертки.

а — общий случай, $M = 2$, $N = 4$, $L = 3$: *б* — импульсный отклик гауссовой формы, $M = 8$, $N = 16$, $L = 9$

Допустим, что импульсный отклик является пространственно-инвариантным и разделимым, т. е.

$$\mathbf{B} = \mathbf{B}_C \otimes \mathbf{B}_R, \tag{9.2.24}$$

где матрицы $\mathbf{B}_R$ и $\mathbf{B}_C$ размера $M\times N$ имеют следующую структуру:

$$\mathbf{B}_R = \begin{bmatrix} h_R(L) & h_R(L-1) & \cdots & h_R(1) & 0 & \cdots & 0 \\ 0 & h_R(L) & \cdots & h_R(2) & h_R(1) & & \vdots \\ \vdots & & \ddots & & & \ddots & 0 \\ 0 & & & 0 & h_R(L) & \cdots & h_R(1) \end{bmatrix}. \tag{9.2.25}$$

В этом случае вычисление двумерной свертки сводится к последовательному выполнению сверток по строкам и столбцам матрицы исходного изображения. Таким образом,

$$\mathbf{G} = \mathbf{B}_C \mathbf{F} \mathbf{B}_R^T. \qquad (9.2.26)$$

Для выполнения операции суперпозиции или вычисления свертки в векторной форме нужно провести M^2L^2 арифметических операций, причем в это число не входят умножения на нулевые элементы матрицы **B**. Если же оператор свертки разделим, то для вычислений в матричной форме достаточно $ML\,(M+N)$ арифметических операций.

Предположим, что при одном и том же импульсном отклике размера $L \times L$ к массиву отсчетов изображения размера $N \times N$ с целью моделирования процесса непрерывной суперпозиции применены оператор суперпозиции конечных массивов и дискретизованный оператор суперпозиции. Тогда массив, обработанный первым оператором, совпадает с массивом, полученным в результате действия второго оператора, который окружен полосой из $L-1$ лишних элементов [1]), и наоборот, если размеры обработанных массивов будут одинаковыми, то $L-1$ пограничных элеметов массива, полученного с помощью оператора конечной суперпозиции, окажутся искаженными. Таким образом, следует с осторожностью применять оператор суперпозиции конечных массивов для моделирования непрерывных процессов.

9.3. ОПЕРАТОР ЦИКЛИЧЕСКОЙ СУПЕРПОЗИЦИИ

При использовании оператора циклической суперпозиции массивы отсчетов входного и выходного изображений, а также импульсного отклика должны быть периодическими по пространственным переменным с одинаковым периодом. Для единства изложения примем, как и прежде, что все эти массивы имеют конечные размеры. Допустим также, что массив исходных отсчетов $F(n_1, n_2)$ размером $N \times N$ помещен в левый верхний угол массива, содержащего $J \times J$ нулей $(J > N)$, т. е. образуется расширенный массив

$$F_E(n_1, n_2) = F(n_1, n_2) \quad \text{при} \quad 1 \leqslant n_i \leqslant N, \qquad (9.3.1а)$$

$$F_E(n_1, n_2) = 0 \quad \text{при} \quad N+1 \leqslant n_i \leqslant J. \qquad (9.3.1б)$$

Аналогично формируется расширенный массив отсчетов импульсного отклика:

$$H_E(l_1, l_2; m_1, m_2) = H(l_1, l_2; m_1, m_2) \quad \text{при} \quad 1 \leqslant l_i \leqslant L, \qquad (9.3.2а)$$

$$H_E(l_1, l_2; m_1, m_2) = 0 \quad \text{при} \quad L+1 \leqslant l_i \leqslant J. \qquad (9.3.2б)$$

[1]) Исключением является случай, когда вдоль границы входного массива лежит полоса из $(L-1)/2$ нулевых элементов. Реальным примером этого может служить смазанный снимок Луны на фоне ночного неба.

Далее образуем периодически продолженные массивы $F_E\ (n_1,\ n_2)$ и $H_E\ (l_1,\ l_2;\ m_1,\ m_2)$, повторяя (размножая) расширенные массивы с периодом в J отсчетов. Циклическая свертка этих массивов по определению равна

$$K_E(m_1,\ m_2) = \sum_{n_1=1}^{J} \sum_{n_2=1}^{J} F_P(n_1,\ n_2)\, H_P(m_1 - n_1 + 1,\ m_2 - n_2 + 1;\ m_1,\ m_2). \qquad (9.3.3)$$

Бросается в глаза сходство данного выражения с равенством (9.1.1), описывающим суперпозицию конечных массивов. В самом деле, если J выбрано так, что $J = N + L - 1$, то $F_E\ (m_1,\ m_2) = F\ (m_1,\ m_2)$ при $1 \leqslant m_i \leqslant M$. Следует также отметить сходство циклической суперпозиции с суперпозицией дискретизованных массивов. Эти соотношения становятся более ясными, когда используется векторное представление циклической суперпозиции.

Допустим, что массивы F_E и K_E представлены соответственно $(J^2 \times 1)$-компонентными векторами $\mathbf{f}_E$ и $\mathbf{k}_E$. Тогда для операции циклической суперпозиции можно записать соотношение

$$\mathbf{k}_E = \mathbf{C}\mathbf{f}_E, \qquad (9.3.4)$$

где $\mathbf{C}$ — матрица элементов массива H_E размера $J^2 \times J^2$. Оператор циклической суперпозиции удобно выразить в виде блочной матрицы с блоками $\mathbf{C}_{m,n}$ размера $J \times J$:

$$\mathbf{C} = \left[\begin{array}{ccccc|ccc}
\mathbf{C}_{1,1} & 0 & 0 & \cdots & 0 & \mathbf{C}_{1,J-L+2} & \cdots & \mathbf{C}_{1,J} \\
\mathbf{C}_{2,1} & \mathbf{C}_{2,2} & 0 & \cdots & 0 & 0 & \cdots & \mathbf{C}_{2,J} \\
\vdots & \vdots & \ddots & & & & \ddots & \vdots \\
 & & & & & & & \mathbf{C}_{L-1,J} \\
\hline
\mathbf{C}_{L,1} & \mathbf{C}_{L,2} & & & & & & 0 \\
0 & \mathbf{C}_{L+1,2} & & \ddots & & & & \vdots \\
0 & & & & & \ddots & & \vdots \\
\vdots & & & & & & & 0 \\
0 & & 0 & & \mathbf{C}_{J,J-L+1} & \mathbf{C}_{J,J-L+2} & \cdots & \mathbf{C}_{J,J}
\end{array}\right], \qquad (9.3.5)$$

где

$$C_{m_2,\, n_2}(m_1,\ n_1) = H_E(k_1,\ k_2;\ m_1,\ m_2), \qquad (9.3.6)$$

причем $1 \leqslant n_i \leqslant J$ и $1 \leqslant m_i \leqslant J$, а $k_i = (m_i - n_i + 1)_{\mathrm{mod}\, J}$ и $H_E\ (0,\ 0) = 0$. Следует отметить, что каждая строка и каждый столбец блочной матрицы $\mathbf{C}$ содержат L ненулевых блоков. Если массив отсчетов импульсного отклика является пространственно-инвариантным, то

$$\mathbf{C}_{m_2,\, n_2} = \mathbf{C}_{m_2+1,\ n_2+1} \qquad (9.3.7)$$

и любую строку (или столбец) можно получить путем циклической перестановки блоков первой строки (или первого столбца). На рис. 9.3.1, *а* приведен пример оператора циклической свертки,

$$
\mathbf{H}=\begin{bmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{bmatrix}
\qquad
\mathbf{C}=\left[\begin{array}{cc|cc|cc|cc|cccc|cccc}
11 & 0 & 31 & 21 & 0 & 0 & 0 & 0 & 13 & 0 & 33 & 23 & 12 & 0 & 32 & 22 \\
21 & 11 & 0 & 31 & 0 & 0 & 0 & 0 & 23 & 13 & 0 & 33 & 22 & 12 & 0 & 32 \\
31 & 21 & 11 & 0 & 0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 \\
0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 \\
\hline
12 & 0 & 32 & 22 & 11 & 0 & 31 & 21 & 0 & 0 & 0 & 0 & 13 & 0 & 33 & 23 \\
22 & 12 & 0 & 32 & 21 & 11 & 0 & 31 & 0 & 0 & 0 & 0 & 23 & 13 & 0 & 33 \\
32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 \\
0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 & 0 & 33 & 23 & 13 \\
\hline
13 & 0 & 33 & 23 & 12 & 0 & 32 & 22 & 11 & 0 & 31 & 21 & 0 & 0 & 0 & 0 \\
23 & 13 & 0 & 33 & 22 & 12 & 0 & 32 & 21 & 11 & 0 & 31 & 0 & 0 & 0 & 0 \\
\hline
33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 & 0 \\
0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 13 & 0 & 33 & 23 & 12 & 0 & 32 & 22 & 11 & 0 & 31 & 21 \\
0 & 0 & 0 & 0 & 23 & 13 & 0 & 33 & 22 & 12 & 0 & 32 & 21 & 11 & 0 & 31 \\
\hline
0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11 & 0 \\
0 & 0 & 0 & 0 & 0 & 33 & 23 & 13 & 0 & 32 & 22 & 12 & 0 & 31 & 21 & 11
\end{array}\right]
$$

а

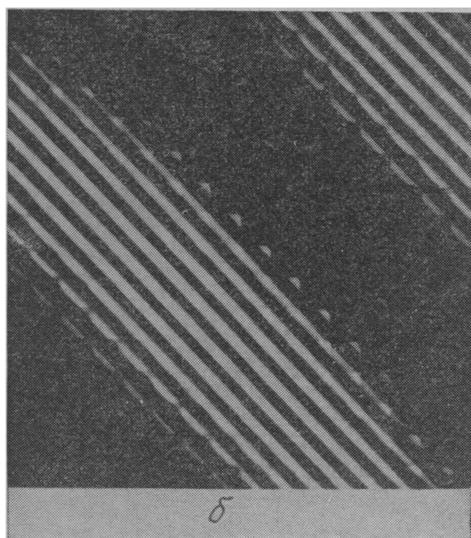


Рис. 9.3.1. Примеры матриц оператора циклической свертки.

а — общий случай, $J = 4$, $L = 3$; *б* — импульсный отклик гауссовой формы, $J = 16$, $L = 9$.

когда входной и выходной массивы имеют размеры 16×16 ($J = 4$), а размер массива отсчетов импульсного отклика равен 3×3 ($L = 3$). На рис. 9.3.1, *б* показана структура матрицы того же оператора при $J = 16$ и $L = 9$, когда импульсный отклик имеет гауссову форму.

Если же импульсный отклик является пространственно-инвариантным и разделимым, то

$$\mathbf{C} = \mathbf{C}_C \otimes \mathbf{C}_R, \tag{9.3.8}$$

где $\mathbf{C}_R$ и $\mathbf{C}_C$ — матрицы вида

$$\mathbf{C}_R = \begin{bmatrix} h_R(1) & 0 & \cdots & 0 & h_R(L) & \cdots & h_R(2) \\ h_R(2) & h_R(1) & \cdots & 0 & 0 & \cdots & h_R(3) \\ \vdots & & & & & & \vdots \\ h_R(L-1) & & & & & & h_R(L) \\ h_R(L) & h_R(L-1) & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & h_R(L) & & & & & \vdots \\ \vdots & & & & & & 0 \\ 0 & \cdots & \cdots 0 & h_R(L) & \cdots & \cdots & h_R(1) \end{bmatrix}, \tag{9.3.9}$$

размер которых равен $J \times J$. При этом двумерная циклическая свертка вычисляется в соответствии с соотношением

$$\mathbf{K}_E = \mathbf{C}_C \dot{\mathbf{E}}_F \mathbf{C}_R^T. \tag{9.3.10}$$

9.4. СООТНОШЕНИЕ МЕЖДУ ЛИНЕЙНЫМИ ОПЕРАТОРАМИ

Из матрицы оператора циклической суперпозиции $\mathbf{C}$ можно получить матрицы оператора суперпозиции конечных массивов $\mathbf{D}$ и дискретизованного оператора суперпозиции $\mathbf{B}$. Для этого следует ввести выделяющие матрицы

$$\mathbf{S1}_J^{(K)} = [\overbrace{\mathbf{I}_K}^{K} \vdots \overbrace{\mathbf{0}}^{J-K}]\}K, \tag{9.4.1а}$$

$$\mathbf{S2}^{(K)} = [\overbrace{\mathbf{0}}^{L-1} \vdots \overbrace{\mathbf{I}_K}^{K} \vdots \overbrace{\mathbf{0}}^{J-K-L+1}]\}K, \tag{9.4.1б}$$

где $\mathbf{I}_K$ — единичная матрица размера $K \times K$. Приведем соотношения, связывающие эти матрицы с матрицами, полученными из них обобщенным обращением и транспонированием:

$$[\mathbf{S1}^{(K)}]^- = [\mathbf{S1}^{(K)}]^T, \tag{9.4.2а}$$

$$[[\mathbf{S1}_J^{(K)}]^T]^- = \mathbf{S1}^{(K)}, \tag{9.4.2б}$$

$$[\mathbf{S2}^{(K)}]^- = [\mathbf{S2}_J^{(K)}]^T, \tag{9.4.2в}$$

$$[[\mathbf{S2}_J^{(K)}]^T]^- = \mathbf{S2}_J^{(K)}. \tag{9.4.2г}$$

Анализируя структуру различных линейных операторов, можно показать, что

$$\mathbf{D} = [\mathbf{S1}_J^{(M)} \otimes \mathbf{S1}_J^{(M)}]\,\mathbf{C}\,[\mathbf{S1}_J^{(N)} \otimes \mathbf{S1}_J^{(N)}]^T, \quad M > N, \tag{9.4.3а}$$

$$\mathbf{B} = [\mathbf{S2}^{(M)} \otimes \mathbf{S2}_J^{(M)}]\,\mathbf{C}\,[\mathbf{S1}_J^{(N)} \otimes \mathbf{S1}_J^{(N)}]^T, \quad N > M. \tag{9.4.3б}$$

Таким образом, матрица $\mathbf{D}$ образуется посредством выделения первых M строк и N столбцов блока $\mathbf{C}_{m,\,n}$ матрицы $\mathbf{C}$. При этом во всех других блоках также выделяются первые M строк и N столбцов. Аналогично из матрицы $\mathbf{C}$ можно образовать матрицу $\mathbf{B}$. Элементы матрицы $\mathbf{C}$, из которых формируются матрицы $\mathbf{D}$ и $\mathbf{B}$, на рис. 9.3.1, *а* заключены в рамки.

Из определения (9.3.1) расширенного массива отсчетов исходного изображения следует, что вектор отсчетов конечного исходного изображения $\mathbf{f}$ образуется из вектора расширенного изображения $\mathbf{f}_E$ при помощи операции выделения:

$$\mathbf{f} = [\mathbf{S1}_J^{(N)} \otimes \mathbf{S1}_J^{(N)}]\,\mathbf{f}_E, \tag{9.4.4а}$$

$$\mathbf{f}_E = [\mathbf{S1}_J^{(N)} \otimes \mathbf{S1}_J^{(N)}]^T\,\mathbf{f}. \tag{9.4.4б}$$

Можно также показать, что выходной вектор оператора суперпозиции конечных массивов можно получить из выходного вектора

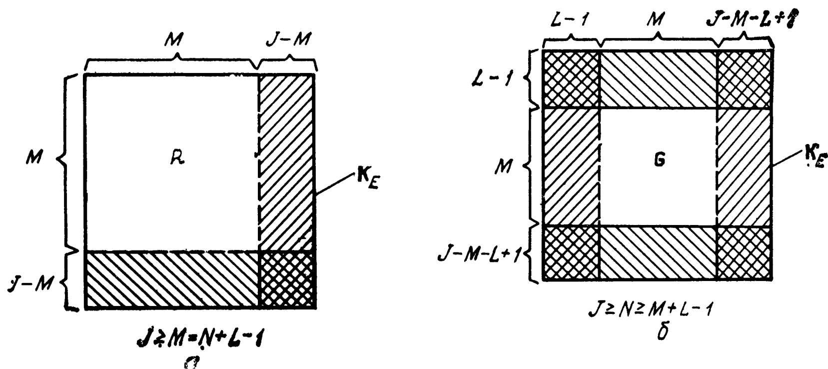


Рис. 9.4.1. Расположение матриц $\mathbf{Q}$ и $\mathbf{G}$ в матрице $\mathbf{K}_E$.
а — оператор суперпозиции конечных массивов; *б* — дискретизованный оператор суперпозиции

оператора циклической суперпозиции с помощью операции выделения:

$$\mathbf{q} = [\mathbf{S1}_J^{(M)} \otimes \mathbf{S1}_J^{(M)}]\,k_E. \tag{9.4.5а}$$

Между векторами существует и обратное соотношение

$$\mathbf{k}_E = \left[\mathbf{S1}_J^{(M)} \otimes \mathbf{S1}_J^{(M)}\right]^T \mathbf{q}. \tag{9.4.5б}$$

Для дискретизованного оператора суперпозиции

$$\mathbf{g} = [\mathbf{S2}_J^{(M)} \otimes \mathbf{S2}_J^{(M)}]\,\mathbf{k}_E, \tag{9.4.6}$$

однако обратный переход от $\mathbf{g}$ к $\mathbf{k}_E$ выполнить не удается в силу недоопределенности дискретизованного оператора суперпозиции. Преобразуя $\mathbf{q}$ и $\mathbf{k}_E$ в матричную форму, из соотношения (9.4.5а) можно получить равенство

$$\mathbf{Q} = \sum_{m=1}^{M} \sum_{n=1}^{J} \mathbf{M}_m^T [\mathbf{S1}_J^{(M)} \otimes \mathbf{S1}_J^{(M)}]\, \mathbf{N}_n \mathbf{K}_E \mathbf{v}_n \mathbf{u}_m^T. \tag{9.4.7}$$

Поскольку оператор выделения обладает свойством разделимости, формула (9.4.7) упрощается к виду

$$\mathbf{Q} = [\mathbf{S1}_J^{(M)}]\, \mathbf{K}_E\, [\mathbf{S1}_J^{(M)}]^T. \tag{9.4.8}$$

Аналогично из равенства (9.4.6), относящегося к дискретизованному оператору суперпозиции, можно получить соотношение

$$\mathbf{G} = [\mathbf{S2}^{(M)}]\, \mathbf{K}_E\, [\mathbf{S2}_J^{(M)}]^T. \tag{9.4.9}$$

На рис. 9.4.1 показано расположение элементов матрицы $\mathbf{K}_E$, из которых формируются матрицы для оператора суперпозиции конечных массивов $\mathbf{Q}$ и для дискретизованного оператора суперпозиции $\mathbf{G}$.

Итак, для обоих операторов выходные векторы можно получить из результата циклической суперпозиции с помощью операции выделения части элементов. Как показано в гл. 11, этот факт позволяет упростить вычисления.

ЛИТЕРАТУРА

1. Pratt W. K., Vector Formulation of Two Dimensional Signal Processing Operations, *J. Computer Graphics and Image Processing*, **4**, 1, 1—24 (March 1975), Academic Press, New York.
2. McCalla T. R., Introduction to Numerical Methods and FORTRAN Programming, Wiley, New York, 1967.
3. Oppenheim A. V., Schaefer R. W., Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N. J., 1975. [Имеется перевод: Оппенгейм А. В., Шафер Р. В., Цифровая обработка сигналов. — М.: Связь, 1979.]
4. Papoulis A., Systems and Transforms with Applications in Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Папулис А., Теория систем и преобразований в оптике. — М.: Мир, 1971.]

Глава 10

ДВУМЕРНЫЕ УНИТАРНЫЕ ПРЕОБРАЗОВАНИЯ

Существуют три основные области применения двумерных унитарных преобразований для обработки изображений. Во-первых, преобразования используются для выделения характерных признаков изображения. Так, например, постоянная составляющая спектра Фурье пропорциональна средней яркости изображения, а высокочастотные составляющие характеризуют величину и ориентацию его контуров. Другой областью применения преобразований является кодирование изображений, когда ширина спектра уменьшается за счет отбрасывания или грубого квантования малых по величине коэффициентов преобразования. Третья область приложений — это сокращение размерности при выполнении вычислений. Иначе говоря, в процессе обработки (например, фильтрации) малые коэффициенты преобразования можно отбросить без заметного ухудшения качества обработки. В последующих главах эти применения будут описаны более детально. В этой главе рассмотрены свойства преобразований, наиболее часто встречающихся при обработке изображений.

10.1. ОПЕРАТОРЫ УНИТАРНЫХ ПРЕОБРАЗОВАНИЙ

Унитарные преобразования являются частным случаем линейных преобразований, когда линейный оператор [см. (8.1.1)] точно обратим, а его ядро удовлетворяет условиям ортогональности [1, 2]. В результате прямого унитарного преобразования матрицы изображения $F(n_1, n_2)$ размера $N_1 \times N_2$ образуется матрица преобразованного изображения того же размера, элементы которой по определению равны

$$\mathcal{F}(m_1, m_2) = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} F(n_1, n_2) A(n_1, n_2; m_1, m_2), \qquad (10.1.1)$$

где $A(n_1, n_2; m_1, m_2)$ — ядро прямого преобразования. Исходное изображение можно получить с помощью обратного преобразования, описываемого соотношением

$$F(n_1, n_2) = \sum_{m_1=1}^{N_1} \sum_{m_2=1}^{N_2} \mathcal{F}(m_1, m_2) B(n_1, n_2; m_1, m_2), \qquad (10.1.2)$$

где $B(n_1, n_2; m_1, m_2)$ — ядро обратного преобразования. Преобразование является унитарным, если выполняются следующие условия ортогональности:

$$\sum_{m_1} \sum_{m_2} A(n_1, n_2; m_1, m_2) A^*(j_1, j_2; m_1, m_2) = \delta(n_1 - j_1, n_2 - j_2), \tag{10.1.3а}$$

$$\sum_{m_1} \sum_{m_2} B(n_1, n_2; m_1, m_2) B^*(j_1, j_2; m_1, m_2) = \delta(n_1 - j_1, n_2 - j_2), \tag{10.1.3б}$$

$$\sum_{n_1} \sum_{n_2} A(n_1, n_2; m_1, m_2) A^*(n_1, n_2; k_1, k_2) = \delta(m_1 - k_1, m_2 - k_2), \tag{10.1.3в}$$

$$\sum_{n_1} \sum_{n_2} (n_1, n_2; m_1, m_2) B^*(n_1, n_2; k_1, k_2) = \delta(m_1 - k_1, m_2 - k_2). \tag{10.1.3г}$$

Преобразование называют разделимым, если оба его ядра можно представить в следующей форме:

$$A(n_1, n_2; m_1, m_2) = A_C(n_1, m_1) A_R(n_2, m_2), \tag{10.1.4а}$$

$$B(n_1, n_2; m_1, m_2) = B_C(n_1, m_1) B_R(n_2, m_2), \tag{10.1.4б}$$

где через A_C (или B_C) и A_R (или B_R) обозначены соответственно одномерные операторы преобразования столбцов и строк. Результат воздействия оператора разделимого двумерного унитарного преобразования можно находить в два этапа. Сначала выполняется одномерное преобразование по всем столбцам матрицы изображения, причем образуется матрица с элементами

$$P(m_1, n_2) = \sum_{n_1=1}^{N_1} F(n_1, n_2) A_C(n_1, m_1). \tag{10.1.5}$$

Затем выполняется второе одномерное преобразование по всем строкам полученной матрицы, в результате которого образуется массив чисел вида

$$\mathcal{F}(m_1, m_2) = \sum_{n_2=1}^{N_2} P(m_1, n_2) A_R(n_2, m_2). \tag{10.1.6}$$

Унитарные преобразования удобно записывать с помощью векторных обозначений [3]. Допустим, что $\mathbf{F}$ и $\mathbf{f}$ — матричное и векторное представления массива отсчетов исходного изображения, а $\mathcal{F}$ и $\mathcal{f}$ — матричное и векторное представления преобразованного изображения. Тогда двумерное унитарное преобразование в векторной форме выражается соотношением

$$\mathcal{f} = \mathbf{A}\mathbf{f}, \tag{10.1.7}$$

где **A** — матрица прямого преобразования. Обратное преобразование записывается как

$$f = \mathbf{B}\mathfrak{f}, \tag{10.1.8}$$

где **B** — матрица обратного преобразования. Очевидно, что

$$\mathbf{B} = \mathbf{A}^{-1}. \tag{10.1.9}$$

Для унитарных преобразований обратная матрица удовлетворяет соотношению

$$\mathbf{A}^{-1} = \mathbf{A}^{*T}. \tag{10.1.10}$$

В этом случае матрицу **A** называют унитарной. Действительная унитарная матрица называется ортогональной матрицей, и для нее справедливо соотношение

$$\mathbf{A}^{-1} = \mathbf{A}^{T}. \tag{10.1.11}$$

Если ядра преобразования разделимы, так что

$$\mathbf{A} = \mathbf{A}_C \otimes \mathbf{A}_R, \tag{10.1.12}$$

где $\mathbf{A}_R$ и $\mathbf{A}_C$ — унитарные матрицы преобразования по строкам и столбцам, то матрицу преобразованного изображения можно получить из матрицы исходного изображения с помощью равенства

$$\mathcal{F} = \mathbf{A}_C \mathbf{F} \mathbf{A}_R^T. \tag{10.1.13а}$$

Обратное преобразование определяется соотношением

$$\mathbf{F} = \mathbf{B}_C \mathcal{F} \mathbf{B}_R^T, \tag{10.1.13б}$$

где $\mathbf{B}_C = \mathbf{A}_C^{-1}$ и $\mathbf{B}_R = \mathbf{A}_R^{-1}$.

Разделимые унитарные преобразования можно также представлять в виде взвешенной суммы матричных произведений вектор-столбцов, сформированных из элементов матриц. Пусть $\mathbf{a}_C(n_1)$ и $\mathbf{a}_R(n_2)$ обозначают n_1-й и n_2-й столбцы матриц $\mathbf{A}_C^T$ и $\mathbf{A}_R^T$. Нетрудно показать, что

$$\mathcal{F} = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} F(n_1, n_2)\, \mathbf{a}_C(n_1)\, \mathbf{a}_R^T(n_2). \tag{10.1.14а}$$

Аналогично

$$\mathbf{F} = \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \mathcal{F}(m_1, m_2)\, \mathbf{b}_C(m_1)\, \mathbf{b}_R^T(m_2), \tag{10.1.14б}$$

где $\mathbf{b}_C(m_1)$ и $\mathbf{b}_R(m_2)$ — m_1-й и m_2-й столбы матриц $\mathbf{B}_C^T$ и $\mathbf{B}_R^T$ соответственно. Матричные произведения векторов, входящие в суммы (10.1.14), образуют последовательности матриц, называемых ба-

зисными матрицами, на основе которых и производится разложение матрицы **F** исходного изображения или преобразованного изображения $\mathcal{F}$.

Возможны различные интерпретации унитарных преобразований. Преобразование изображения можно рассматривать как разложение исходного изображения в обобщенный двумерный спектр [4]. Каждая спектральная составляющая характеризует вклад соответствующей спектральной (базисной) функции в энергию исходного изображения. При такой трактовке понятие частоты можно обобщить так, чтобы оно было применимо не только к синусам и косинусам, но и к другим функциям, на которых основываются преобразования. Подобный обобщенный спектральный анализ полезен для изучения тех конкретных разложений, которые в наибольшей мере подходят для данного класса изображений. Наглядное представление о преобразованиях изображений можно получить и по-другому, рассматривая преобразование как поворот многомерной системы координат. Одним из главных свойств унитарного преобразования является сохранение метрики. Например, евклидово расстояние между двумя изображениями равно евклидову расстоянию между их образами. Третья возможность интерпретации преобразований заключается в том, что равенство (10.1.2) можно рассматривать как способ составления изображения из набора двумерных функций $B(n_1, n_2; m_1, m_2)$, каждая из которых соответствует определенной точке (m_1, m_2) плоскости обобщенных частот. В подобной интерпретации ядро $B(n_1, n_2; m_1, m_2)$ называют двумерной базисной функцией, а коэффициент $\mathcal{F}(m_1, m_2)$ указывает «вес» этой базисной функции, необходимый для получения рассматриваемого изображения.

Для упрощения анализа свойств двумерных унитарных преобразований ниже в этой главе будет принято, что все массивы являются квадратными и имеют размер $N \times N$. Кроме того, будут изменены обозначения и нумерация всех индексов в выражениях вида (10.1.1) и (10.1.2). Таким образом, массив отсчетов исходного изображения будет обозначен через $F(j, k)$, где $j, k = 0, 1, \ldots, N-1$, а массив коэффициентов преобразования — через $\mathcal{F}(u, v)$, где $u, v = 0, 1, \ldots, N-1$. С учетом этого прямое унитарное преобразование записывается в следующем виде:

$$\mathcal{F}(u, v) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) A(j, k; u, v), \qquad (10.1.15а)$$

а обратное преобразование — как

$$F(j, k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) B(j, k; u, v). \qquad (10.1.15б)$$

10.2. ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Дискретное двумерное преобразование Фурье матрицы отсчетов изображения определяется [5—10] в виде ряда

$$\mathcal{F}(u, v) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) \exp\left\{-\frac{2\pi i}{N}(uj + vk)\right\}, \tag{10.2.1а}$$

где $i = \sqrt{-1}$, а дискретное обратное преобразование имеет вид

$$F(j, k) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) \exp\left\{\frac{2\pi i}{N}(uj + vk)\right\}. \tag{10.2.1б}$$

По аналогии с терминологией непрерывного преобразования Фурье переменные u, v называют пространственными частотами. Следует отметить, что не все исследователи пользуются определением (10.2.1); одни предпочитают помещать все масштабные постоянные в выражение для обратного преобразования, а другие изменяют знаки в ядрах на противоположные.

Поскольку ядра преобразования симметричны и разделимы, двумерное преобразование можно выполнить в виде последовательных одномерных преобразований по строкам и столбцам матрицы изображения. Базисными функциями преобразования являются экспоненты с комплексными показателями, которые можно разложить на синусную и косинусную составляющие. Таким образом,

$$A(j, k; u, v) = \exp\left\{-\frac{2\pi i}{N}(uj + vk)\right\} = \cos\left\{\frac{2\pi}{N}(uj + vk)\right\} - i\sin\left\{\frac{2\pi}{N}(uj + vk)\right\}, \tag{10.2.2а}$$

$$B(j, k; u, v) = \exp\left\{\frac{2\pi i}{N}(uj + vk)\right\} = \cos\left\{\frac{2\pi}{N}(uj + vk)\right\} + i\sin\left\{\frac{2\pi}{N}(uj + vk)\right\}. \tag{10.2.2б}$$

На рис. 10.2.1 приведены графики синусных и косинусных составляющих одномерных базисных функций преобразования Фурье для $N = 16$. Видно, что для низких частот эти функции являются грубыми аппроксимациями непрерывных синусоид. С повышением частоты сходство базисных функций с синусоидами теряется. Для наивысшей частоты базисная функция представляет собой меандр. Можно заметить также избыточность наборов синусных и косинусных составляющих.

Спектр изображения имеет много интересных структурных особенностей. Спектральная составляющая в начале координат частотной плоскости

$$\mathscr{F}(0,\ 0) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j,\ k) \qquad (10.2.3)$$

равна увеличенному в N раз среднему (по исходной плоскости) значению яркости изображения. Подставив в равенство (10.2.1а)

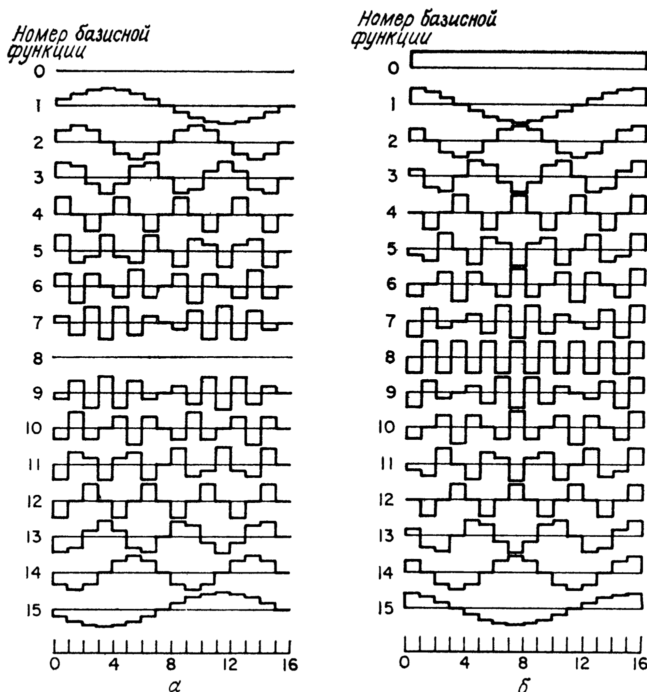


Рис. 10.2.1. Базисные функции преобразования Фурье.
а — синусная (нечетная) составляющая; *б* — косинусная (четная) составляющая.

$u = u + mN$ и $v = v + nN$, где m и n — постоянные, получим

$$\mathscr{F}(u + mN,\ v + nN) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j,\ k) \exp\left\{-\frac{2\pi i}{N}(uj + vk)\right\} \times$$

$$\times \exp\{-2\pi i (mj + nk)\}. \qquad (10.2.4)$$

При любых целочисленных значениях m и n второй экспоненциальный множитель равенства (10.2.4) превращается в единицу. Таким образом, при $m, n = 0,\ \pm 1,\ \pm 2, \dots$

$$\mathscr{F}(u + mN,\ v + nN) = \mathscr{F}(u,\ v), \qquad (10.2.5$$

что свидетельствует о периодичности частотной плоскости. Рис. 10.2.2, *а* иллюстрирует этот результат.

Двумерный спектр Фурье изображения является по существу представлением двумерного поля в виде ряда Фурье. Для того чтобы такое представление было справедливым, исходное изображение также должно обладать периодической структурой, т. е. (как показано на рис. 10.2.2, *б*) иметь рисунок, повторяющийся по вертикали и горизонтали. Таким образом, правый край изобра-

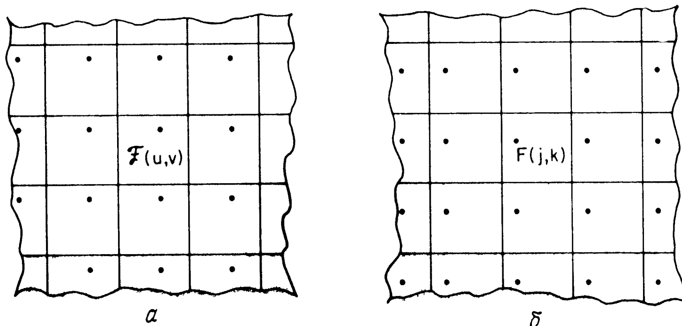


Рис. 10.2.2. Периодическое продолжение изображения и спектра Фурье.
а — спектр; *б* — исходное изображение

жения примыкает к левому, а верхний край — к нижнему. Из-за разрывов значений яркости в этих местах в спектре изображения возникают дополнительные составляющие, лежащие на координатных осях частотной плоскости. Эти составляющие не связаны со значениями яркости внутренних точек изображения, но они необходимы для воспроизведения его резких границ.

Если массив отсчетов изображения описывает поле яркости, то числа $F(j, k)$ будут действительными и положительными. Однако спектр Фурье этого изображения в общем случае имеет комплексные значения. Поскольку спектр содержит $2N^2$ компонент, представляющих действительную и мнимую части или фазу и модуль спектральных составляющих для каждой частоты, может показаться, что преобразование Фурье увеличивает размерность изображения. Это, однако, не так, поскольку $\mathcal{F}(u, v)$ обладает симметрией относительно комплексного сопряжения. Если в равенстве (10.2.4) положить m и n равными целым числам, то после комплексного сопряжения получится равенство

$$\mathcal{F}^*(u+mN,\ v+nN) = \frac{1}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} F(j,\ k)\exp\left\{\frac{2\pi i}{N}(uj+vk)\right\}. \qquad (10.2.6$$

С помощью подстановки $u = -u$ и $v = -v$ можно показать, что

$$\mathcal{F}(u,\ v) = \mathcal{F}^*(-u+mN,\ -v+nN) \qquad (10.2.7)$$

при $m, n = 0, \pm 1, \pm 2, \ldots$. Из-за наличия комплексно-сопряженной симметрии почти половина спектральных составляющих оказывается избыточной, т. е. их можно сформировать из остальных составляющих. На рис. 10.2.3 области расположения избыточных составляющих спектра на частотной плоскости заштрихованы. Избыточными составляющими можно, конечно, считать гармоники, попадающие не в нижнюю, а в правую полуплоскость.

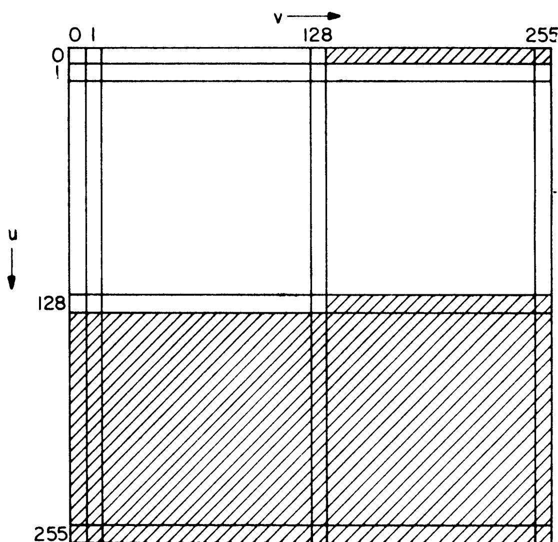


Рис. 10.2.3. Частотная плоскость.

На рис. 10.2.4 приведены фотографии исходного изображения и различных вариантов его фурье-спектра. Поскольку динамический диапазон составляющих спектра гораздо шире, чем линейный участок интервала экспозиций фотопленки, то спектры необходимо подвергнуть сжатию. Сжатие динамического диапазона можно осуществить путем ограничения больших спектральных составляющих или логарифмическим преобразованием всех составляющих спектра согласно соотношению

$$\mathcal{D}(u, v) = \log [a + b \,|\, \mathcal{F}(u, v)|], \tag{10.2.8}$$

где a и b — масштабные постоянные. На рис. 10.2.4, *б* приведены (в логарифмическом масштабе с $a = 1$ и $b = 1$) модули гармоник спектра, вычисленных в соответствии с равенством (10.2.1а). На рис. 10.2.4, *в* показан тот же спектр, в котором наибольшие гармоники, составляющие 25 % всех компонент спектра, были ограничены по величине. При математическом анализе непрерывных сигналов начало координат частотной плоскости обычно

помещают в ее геометрическом центре. Аналогично в дифракционной картине, полученной с помощью когерентной оптической системы для диапозитива с коэффициентом пропускания $F(x, y)$,

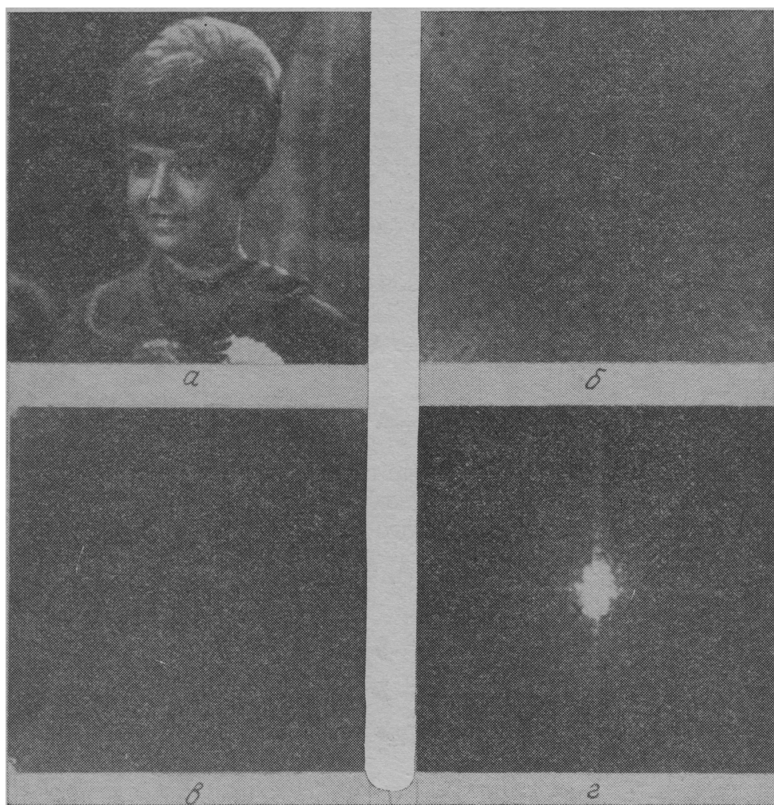


Рис. 10.2.4. Преобразование Фурье изображения «Портрет».

а — исходное изображение; *б* — модули гармоник спектра (в логарифмическом масштабе, без перестановки гармоник); *в* — спектр с ограниченными наибольшими гармониками (без их перестановки); *г* — спектр с ограниченными наибольшими гармониками (после их перестановки).

гармоника с нулевой частотой оказывается в центре. Двумерный дискретный фурье-спектр, найденный с помощью вычислительной машины, путем простой перестановки коэффициентов можно изменить так, чтобы начало координат также оказалось в центре массива. Тот же результат можно получить по-другому, если отсчеты изображения предварительно умножить на коэффициенты вида $(-1)^{j+k}$. Тогда квадранты спектра, рассчитанного согласно формуле (10.2.1a), в процессе вычисления автоматически меняются

местами. Данное утверждение можно доказать с помощью соотношения (10.2.4), положив в нем $m = n = {}^1/_2$. Тогда в силу тождества

$$\exp\{i\pi(j+k)\} = (-1)^{j+k} \qquad (10.2.9)$$

выражение (10.2.4) можно записать в следующем виде:

$$\mathscr{F}(u+N/2,\ v+N/2) = \\ = \frac{1}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} F(j,\ k)(-1)^{j+k}\exp\left\{-\frac{2\pi i}{N}(uj+vk)\right\}. \qquad (10.2.10)$$

На рис. 10.2.4, *г* представлен спектр с переставленными гармониками.

Прямое и обратное преобразования Фурье (10.2.1) можно представить в векторной форме

$$\mathbf{f} = \mathbf{A}\mathbf{f}, \qquad (10.2.11а)$$

а

$$\mathbf{f} = \mathbf{A}^{*T}\mathbf{f}, \qquad (10.2.11б)$$

где f и f — векторы, полученные соответственно путем развертки по столбцам матриц $\mathbf{F}$ и $\mathscr{F}$. Поскольку матрицу преобразования $\mathbf{A}$ можно записать в виде прямого произведения

$$\mathbf{A} = \mathbf{A}_C \otimes \mathbf{A}_R, \qquad (10.2.12)$$

где

$$\mathbf{A}_R = \mathbf{A}_C = \frac{1}{\sqrt{N}}\begin{bmatrix} \mathscr{W}^0 & \mathscr{W}^0 & \mathscr{W}^0 & \dots & \mathscr{W}^0 \\ \mathscr{W}^0 & \mathscr{W}^1 & \mathscr{W}^2 & \dots & \mathscr{W}^{N-1} \\ \mathscr{W}^0 & \mathscr{W}^2 & \mathscr{W}^4 & \dots & \mathscr{W}^{2(N-1)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathscr{W}^0 & \dots & \dots & \dots & \mathscr{W}^{(N-1)^2} \end{bmatrix}, \qquad (10.2.13)$$

$\mathscr{W} = \exp\{-2\pi i/N\}$, то матрицы изображения и спектра связаны соотношениями

$$\mathscr{F} = \mathbf{A}_C\mathbf{F}\mathbf{A}_R \qquad (10.2.14а)$$

и

$$\mathbf{F} = \mathbf{A}_C^*\mathscr{F}\mathbf{A}_R^*. \qquad (10.2.14б)$$

Понятно, что свойства преобразования Фурье, установленные для рядов (10.2.1), сохраняются и при его матричной записи.

Хотя преобразование Фурье имеет много полезных для анализа свойств, у него есть и два существенных недостатка: во-первых, все вычисления приходится производить не с действительными, а с комплексными числами и, во-вторых, ряды сходятся медленно. Последнее замечание, весьма существенное для

задач кодирования изображений, можно пояснить, переписав определение (10.2.1б) в следующей форме:

$$F(j, k) = \lim_{\substack{U_T \to (N+1)/2 \\ V_T \to (N+1)/2}} \left[\underbrace{\frac{1}{N} \sum_u \sum_v}_{u,\, v \in S(U_T,\, V_T)} \mathcal{F}(u, v) \exp\left\{\frac{2\pi i}{N}(uj + vk)\right\} \right], \tag{10.2.15}$$

где $S(U_T, V_T)$ — низкочастотная область частотной плоскости, которая на рис. 10.2.5 заштрихована. Если верхние границы U_T и V_T зафиксированы, а размер изображения N сравнительно невелик, то заключенная в квадратные скобки величина может сильно отличаться от $F(j, k)$, пока U_T и V_T не станут достаточно

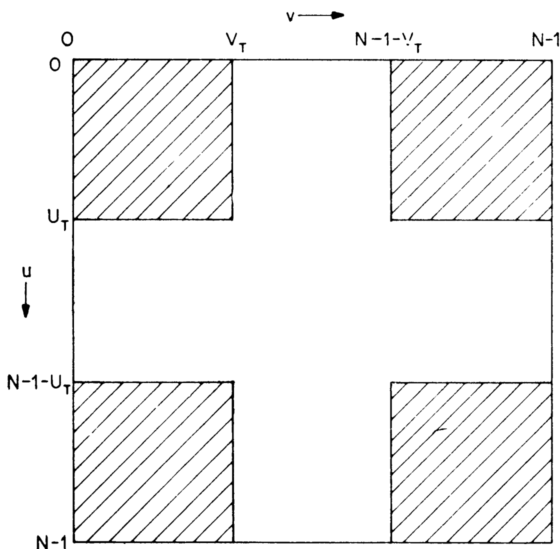


Рис. 10.2.5. Низкочастотная область частотной плоскости.

большими. В дальнейшем вопрос о сходимости преобразований будет проанализирован с количественных позиций. Качественно же можно указать, что плохая сходимость преобразования Фурье обусловлена скачками изображения, возникающими на линиях перехода от левого края изображения к правому и от верхнего к нижнему. Эти разрывы приводят к появлению в спектре больших составляющих с высокими пространственными частотами.

10.3. КОСИНУСНЫЕ ПРЕОБРАЗОВАНИЯ

Известно, что ряд Фурье для любой непрерывной действительной и симметричной (четной) функции содержит только действительные коэффициенты, соответствующие косинусным членам

ряда. В соответствующей интерпретации этот результат можно распространить и на дискретное преобразование Фурье изображений. Существуют [11] два способа получения симметричных изображений (рис. 10.3.1). Согласно первому из них, к изображению вплотную пристраивают его зеркальные отражения. По вто-

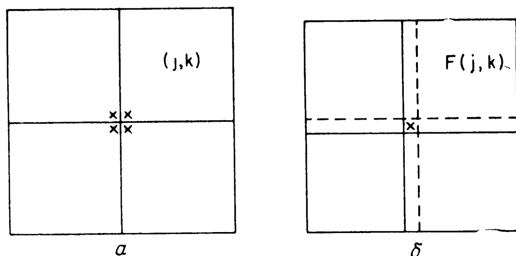


Рис. 10.3.1. Построение симметричного изображения, предназначенного для косинусного преобразования.
а — отражение относительно края; *б* — отражение относительно крайних элементов.

рому методу оригинал и отражения пристраивают, налагая крайние элементы. Таким образом, из первоначального массива, содержащего $N \times N$ элементов, в первом случае (называемом четным косинусным преобразованием) получается массив из $2N \times 2N$ элементов, а во втором случае (называемом нечетным косинусным преобразованием) — массив из $(2N-1) \times (2N-1)$ элементов.

Четное симметричное косинусное преобразование

Допустим, что симметричный массив образован путем зеркального отражения исходного массива относительно его краев согласно соотношению

$$F_s(j, k) = \begin{cases} F(j, k) \text{ при } j \geqslant 0,\ k \geqslant 0, \\ F(-1-j, k) \text{ при } j < 0,\ k \geqslant 0, \\ F(j, -1-k) \text{ при } j \geqslant 0,\ k < 0, \\ F(-1-j, -1-k) \text{ при } j < 0,\ k < 0. \end{cases} \quad (10.3.1)$$

Построенный таким образом массив $F_s(j, k)$ симметричен относительно точки $j = -^1/_2$, $k = -^1/_2$. Вычислив преобразование Фурье для случая, когда начало координат находится в центре симметрии, получим

$$\mathcal{F}_s(u, v) = \frac{1}{2N} \sum_{j=-N}^{N-1} \sum_{k=-N}^{N-1} F_s(j, k) \exp\left\{-\frac{2\pi i}{2N}\left[u\left(j+\frac{1}{2}\right)+v\left(k+\frac{1}{2}\right)\right]\right\}, \quad (10.3.2)$$

где $u, v = -N, \ldots, -1, 0, 1, \ldots, N-1$. Поскольку массив $F_s(j, k)$ симметричен и состоит из действительных чисел, соотношение (10.3.2) можно свести к виду

$$\mathcal{F}_s(u, v) = \\ = \frac{2}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) \cos\left[\frac{\pi}{N} u \left(j + \frac{1}{2}\right)\right] \cos\left[\frac{\pi}{N} v \left(k + \frac{1}{2}\right)\right]. \tag{10.3.3}$$

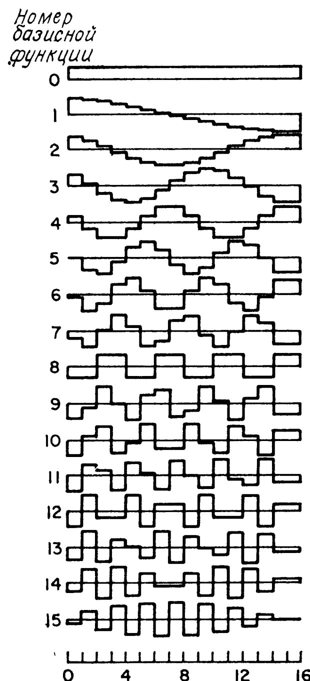


Рис. 10.3.2. Базисные функции косинусного преобразования.

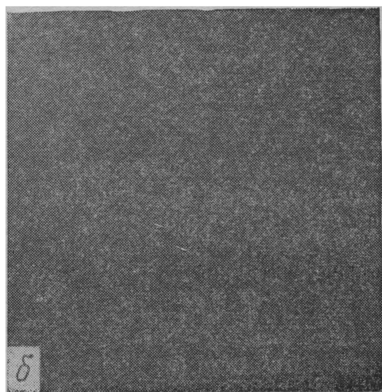
С другой стороны, спектральные составляющие вида (10.3.3) можно найти, вычисляя преобразование Фурье массива $F(j, k)$ по $2N$ точкам:

$$\mathcal{F}_s(u, v) = \\ = \frac{2}{N} \operatorname{Re}\left\{\exp\left[\frac{i\pi u}{2N}\right] \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) \exp\left\{-\frac{2\pi i}{2N}(uj + vk)\right\}\right\}. \tag{10.3.4}$$

Прямое четное косинусное преобразование по определению [12] равно сумме (10.3.3), умноженной на нормирующий мно-

житель, т. е.

$$\mathcal{F}(u, v) = \frac{2}{N} C(u) C(v) \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) \times$$
$$\times \cos\left[\frac{\pi}{N} u \left(j + \frac{1}{2}\right)\right] \cos\left[\frac{\pi}{N} v \left(k + \frac{1}{2}\right)\right], \qquad (10.3.5а)$$



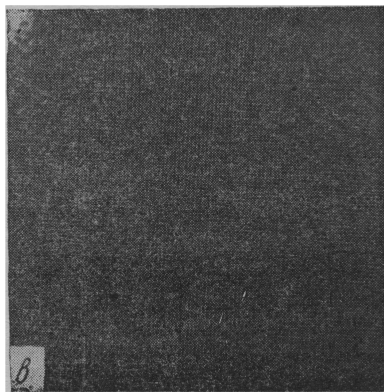


Рис. 10.3.3. Косинусное преобразование изображения «Портрет».
а — исходное изображение; *б* — косинусный спектр в логарифмическом масштабе по оси амплитуд; *в* — спектр с ограниченными наибольшими гармониками.

а обратное преобразование определяется соотношением

$$F(j, k) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) \mathcal{F}(u, v) \times$$
$$\times \cos\left[\frac{\pi}{N} u \left(j + \frac{1}{2}\right)\right] \cos\left[\frac{\pi}{N} v \left(k + \frac{1}{2}\right)\right], \qquad (10.3.5б)$$

где $C(0) = (2)^{-1/2}$, а $C(w) = 1$ при $w = 1, 2, \ldots, N-1$. Оказалось, что базисные функции четного косинусного преобразования принадлежат к классу дискретных полиномов Чебышева [12].

На рис. 10.3.2 приведены графики базисных функций четного симметричного косинусного преобразования при $N = 16$. Образцы спектра, получаемого при четном симметричном косинусном преобразовании, приведены на рис. 10.3.3. Начало координат расположено в верхнем левом углу каждого снимка, что согласуется с принятым в теории матриц порядком. Следует отметить, что здесь, как и в случае преобразования Фурье, основная часть энергии изображения концентрируется в области низких пространственных частот.

Нечетное симметричное косинусное преобразование

При нечетном косинусном преобразовании строение симметричного массива определяется следующим образом:

$$F_s(j, k) = \begin{cases} F(j, k) \text{ при } j, k \geqslant 0, \\ F(-j, k) \text{ при } j < 0, k \geqslant 0, \\ F(j, -k) \text{ при } j \geqslant 0, k < 0, \\ F(-j, -k) \text{ при } j, k < 0. \end{cases} \qquad (10.3.6)$$

Вычисление двумерного преобразования Фурье от такого массива дает

$$\mathcal{F}_s(u, v) = \frac{1}{2N-1} \sum_{j=-N+1}^{N-1} \sum_{k=-N+1}^{N-1} F_s(j, k) \times \times \exp\left\{-\frac{2\pi i}{2N-1}(ju + kv)\right\}, \qquad (10.3.7)$$

где $u, v = -N+1, \ldots, -1, 0, 1, \ldots, N-1$. Поскольку преобразование Фурье обладает свойством симметрии относительно комплексного сопряжения, то для реальных изображений

$$\mathcal{F}_s(u, v) = \mathcal{F}_s^*(-u, -v). \qquad (10.3.8)$$

Следовательно, $\mathcal{F}_s(u, v)$ достаточно вычислять только при неотрицательных значениях индексов (u, v). Кроме того, поскольку функция $F_s(j, k)$ принимает действительные значения и симметрична, то $\mathcal{F}_s(u, v)$ также имеет действительные значения. Таким образом, соотношение (10.3.2) можно представить в следующем виде:

$$\mathcal{F}_s(u, v) = \frac{4}{2N-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \tilde{F}(j, k) \cos\left[\frac{2\pi}{2N-1} ju\right] \cos\left[\frac{2\pi}{2N-1} kv\right], \qquad (10.3.9)$$

где массив $\tilde{F}(j, k)$ получается из матрицы изображения $F(j, k)$ взвешиванием ее элементов в соответствии с формулой

$$\tilde{F}_s(j, k) = \begin{cases} {}^1/_4 F(j, k) \text{ при } j = 0,\ k = 0, \\ {}^1/_2 F(j, k) \text{ при } j = 0,\ k \neq 0, \\ {}^1/_2 F(j, k) \text{ при } j \neq 0,\ k = 0, \\ F(j, k) \quad \text{ при } j \neq 0,\ k \neq 0. \end{cases} \qquad (10.3.10)$$

Нечетное косинусное преобразование есть просто нормированный вариант равенства (10.3.9), причем нормирование проведено так, чтобы базисные функции стали ортонормальными. Таким образом, нечетное косинусное преобразование определяется соотношениями

$$\mathcal{F}(u, v) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \tilde{F}(j, k) \text{ при } u = 0,\ v = 0, \qquad (10.3.11а)$$

$$\mathcal{F}(u, v) = \frac{2}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \tilde{F}(j, k) \cos\left[\frac{2\pi}{2N-1} ju\right] \cos\left[\frac{2\pi}{2N-1} kv\right]$$

$$\text{при } u, v \neq 0. \qquad (10.3.11б)$$

Такое же преобразование

$$\tilde{F}(j, k) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) \text{ при } j = 0,\ k = 0, \qquad (10.3.12а)$$

$$\tilde{F}(j, k) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) \cos\left[\frac{2\pi}{2N-1} ju\right] \cos\left[\frac{2\pi}{2N-1} kv\right]$$

$$\text{при } j, k \neq 0 \qquad (10.3.12б)$$

дает матрицу взвешенных отсчетов $\tilde{F}(j, k)$. Тогда исходный массив можно восстановить с помощью формулы

$$F(j, k) = \begin{cases} 4\tilde{F}(j, k) \text{ при } j = 0,\ k = 0, \\ 2\tilde{F}(j, k) \text{ при } j = 0,\ k \neq 0, \\ 2\tilde{F}(j, k) \text{ при } j \neq 0,\ k = 0, \\ \tilde{F}(j, k) \text{ при } j \neq 0,\ k \neq 0. \end{cases} \qquad (10.3.13)$$

Базисные функции нечетного косинусного преобразования являются разделимыми, так что двумерное нечетное косинусное преобразование можно выполнить с помощью последовательных одномерных преобразований. Кроме того, нечетное косинусное

преобразование можно найти с помощью алгоритма преобразования Фурье при нечетном числе элементов, поскольку

$$\mathcal{F}_s(u, v) = \frac{4}{2N-1} Re \left\{ \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \tilde{F}(j, k) \exp\left\{ -\frac{2\pi i}{2N-1} (ju + kv) \right\} \right\}. \quad (10.3.14)$$

10.4. СИНУСНОЕ ПРЕОБРАЗОВАНИЕ

Быстрое синусное преобразование, предложенное Джейном [13] в качестве аппроксимации преобразования Карунена—Лоэва для марковского процесса, в одномерном случае определяется с помощью базисных функций вида

$$A(u, j) = \sqrt{\frac{2}{N+1}} \sin\left[\frac{(j+1)(u+1)\pi}{N+1}\right], \quad (10.4.1)$$

где $u, j = 0, 1, 2, \ldots, N-1$. Рассмотрим матрицу, содержащую ненулевые элементы только на главной диагонали и двух прилегающих к ней поддиагоналях (так называемую трехдиагональную матрицу):

$$\mathbf{Q} = \begin{bmatrix} 1 & -\alpha & 0 & \ldots & & 0 \\ -\alpha & 1 & -\alpha & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & -\alpha & 1 & -\alpha \\ 0 & \ldots & \ldots & 0 & -\alpha & 1 \end{bmatrix}, \quad (10.4.2)$$

где $\alpha = \rho/(1+\rho^2)$, причем $0 \leqslant \rho \leqslant 1$ — коэффициент корреляции соседних элементов марковского процесса. Можно показать [14], что с помощью унитарной матрицы $\mathbf{A}$, в качестве элементов которой взяты базисные функции (10.4.1), матрицу $\mathbf{Q}$ можно привести к диагональному виду [1]) в том смысле, что

$$\mathbf{AQA}^T = \mathbf{D}, \quad (10.4.3)$$

где $\mathbf{D}$ — диагональная матрица, составленная из элементов

$$D(k, k) = \frac{1-\rho^2}{1 - 2\rho\cos[k\pi/(N+1)] + \rho^2} \quad (10.4.4)$$

при $k = 1, 2, \ldots, N$.

---

[1]) Джейн [15] показал, что косинусное и синусное преобразования обладают общим свойством — с их помощью удается приводить к диагональному виду трехдиагональные матрицы.

Двумерное синусное преобразование определяется соотношением

$$\mathscr{F}(u, v)=\frac{2}{N+1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} F(j, k)\sin\left[\frac{(j+1)(u+1)\pi}{N+1}\right]\times$$
$$\times\sin\left[\frac{(k+1)(v+1)\pi}{N+1}\right], \qquad (10.4.5)$$

а обратное преобразование имеет тот же вид. Синусное преобразование можно вычислить с помощью алгоритма преобразования

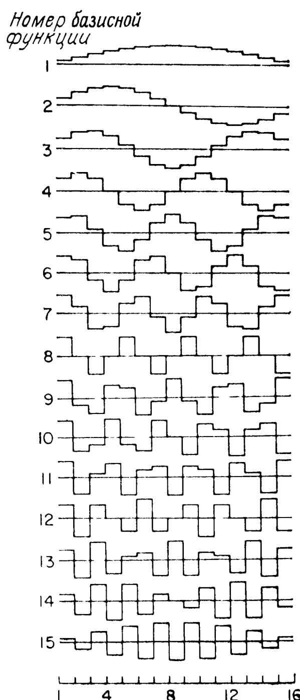


Рис. 10.4.1. Базисные функции синусного преобразования при $N = 15$.

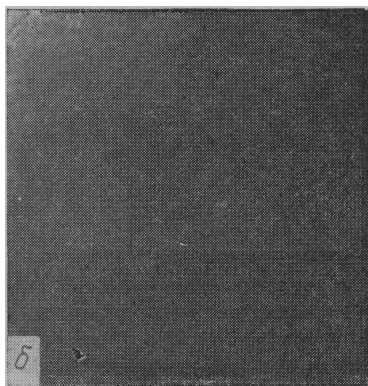
Фурье. Допустим, что массив $\tilde{F}(m, n)$ размером $(2N+2)\times(2N+2)$ образуется в соответствии с равенствами

$$\tilde{F}(m+1, n+1) = F(m, n) \text{ при } 0 \leqslant m \leqslant N-1,\ 0 \leqslant n \leqslant N-1, \qquad (10.4.6а)$$

$$\tilde{F}(m, n) = 0 \text{ в других случаях} \qquad (10.4.6б)$$

Тогда, выделяя мнимую часть коэффициентов Фурье массива $\tilde{F}(m, n)$, можно найти синусное преобразование в виде

$$\mathcal{F}(u, v) = \frac{1}{(N+1)} \sum_{m=0}^{2N+1} \sum_{n=0}^{2N+1} \tilde{F}(m, n) \sin\left[\frac{2(u+1)m}{2N+2}\right] \times \\ \times \sin\left[\frac{2(v+1)n}{2N+2}\right]. \tag{10.4.7}$$



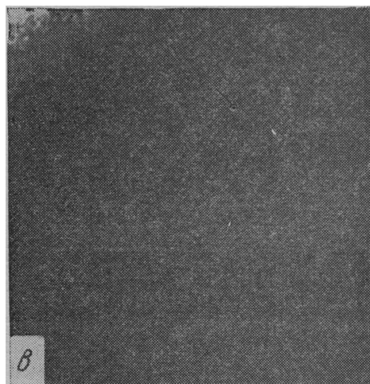


Рис. 10.4.2. Синусное преобразование изображения «Портрет».

а — исходное изображение; *б* — синусный спектр в логарифмическом масштабе по оси амплитуд· *в* — спектр с ограниченными наибольшими гармониками

Графики базисных функций синусного преобразования при $N = 15$ представлены на рис. 10.4.1, а на рис. 10.4.2 приведены фотографии, полученные синусным преобразованием изображения.

10.5. ПРЕОБРАЗОВАНИЕ АДАМАРА

Преобразование Адамара [16, 17] основано на квадратной матрице Адамара [18], элементы которой равны плюс или минус единице, а строки и столбцы образуют ортогональные векторы. Нормированная матрица Адамара N-го порядка удовлетворяет соотношению

$$\mathbf{H}_N \mathbf{H}_N^T = \mathbf{I}_N. \tag{10.5.1}$$

Среди ортонормальных матриц Адамара наименьшей является матрица второго порядка

$$\mathbf{H}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{10.5.2}$$

Известно, что если матрица Адамара порядка N (где $N > 2$) существует, то N делится на 4 без остатка [19]. Пока не удалось определить, существуют ли матрицы Адамара для произвольных N, удовлетворяющих этому условию, однако почти для всех

$$\mathbf{H}_4 = \frac{1}{2} \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array}\right] \qquad \begin{array}{c} \text{Число изменений знака} \\ 0 \\ 3 \\ 1 \\ 2 \end{array}$$

$$\mathbf{H}_8 = \frac{1}{2\sqrt{2}} \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ \hline 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{array}\right] \qquad \begin{array}{c} \text{Число изменений знака} \\ 0 \\ 7 \\ 3 \\ 4 \\ 1 \\ 6 \\ 2 \\ 5 \end{array}$$

Рис. 10.5.1. Неупорядоченные матрицы Адамара четвертого и восьмого порядка.

допустимых N, доходящих до 200, найдены правила построения соответствующих матриц. Наиболее просто удается построить такие матрицы при $N = 2^n$, где n — целое. Если $\mathbf{H}_N$ — матрица Адамара N-го порядка, то матрица

$$\mathbf{H}_{2N} = \frac{1}{\sqrt{2}} \left[\begin{array}{c|c} \mathbf{H}_N & \mathbf{H}_N \\ \hline \mathbf{H}_N & -\mathbf{H}_N \end{array}\right] \tag{10.5.3}$$

также является матрицей Адамара, но порядка $2N$. На рис. 10.5.1 приведены матрицы Адамара четвертого и восьмого порядка, построенные с помощью соотношения (10.5.3).

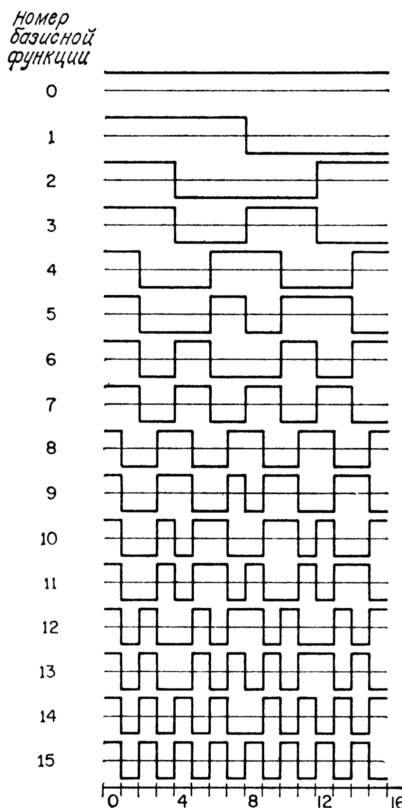


Рис. 10.5.2. Базисные функции преобразования Адамара при $N = 16$.

Хармут [20] предложил частотную интерпретацию матриц Адамара, имеющих блочную структуру (10.5.3). Число изменений знака вдоль каждой строки матрицы Адамара, деленное на два, называется секвентой строки. Можно построить матрицу Адамара порядка $N = 2^n$, в которой число изменений знака в строках принимает значения от 0 до $N - 1$. Унитарные матрицы с такими характеристиками называются матрицами с секвентным свойством.

Строки матрицы Адамара, описываемой соотношением (10.5.3), можно рассматривать как последовательность отсчетов прямоугольных периодических колебаний (сигналов), период которых кратен $1/N$. Подобные непрерывные функции, называемые функ-

циями Уолша[1]) [21], связаны с импульсными функциями Радемахера [22]. Следовательно, матрица Адамара описывает преобразование, связанное с разложением функций по семейству прямо-

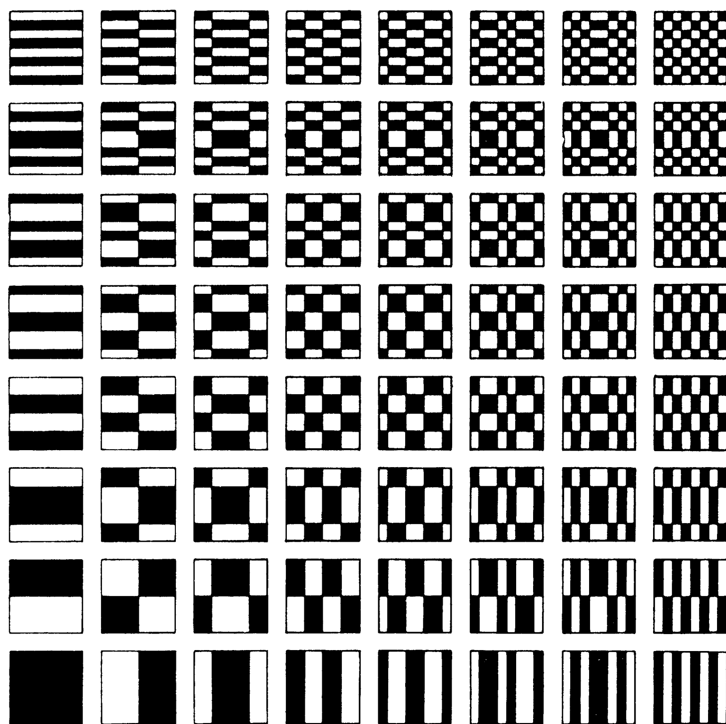


Рис. 10.5.3. Базисные изображения преобразования Адамара при $N = 8$. Черный цвет соответствует значению $+1$, белый — значению -1.

угольных базисных функций, а не по синусам и косинусам, характерным для преобразования Фурье.

Для симметричных матриц Адамара порядка $N = 2^n$ двумерное преобразование Адамара можно представить в виде ряда

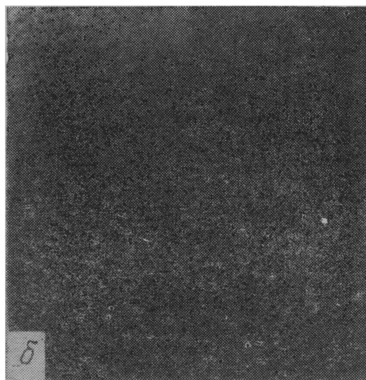
$$\mathcal{F}(u, v) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k)(-1)^{p(j, k, u, v)}, \qquad (10.5.4)$$

[1]) Преобразование Адамара, определяемое с помощью матриц вида (10.5.3), известно также под названием преобразования Уолша. Следует, однако, отметить, что существуют матрицы Адамара (например, матрица 12-го порядка), строки которых не состоят из отсчетов функций Уолша.

где

$$p(j,\ k,\ u,\ v)=\sum_{i=0}^{n-1}(u_i j_i+v_i k_i). \qquad (10.5.5)$$

Переменные u_i, v_i, j_i и k_i равны цифрам в двоичном представлении чисел u, v, j и k соответственно. Так, например, если $u = 13$, то $u_3 = 1$, $u_2 = 1$, $u_1 = 0$ и $u_0 = 1$. Если матрица Адамара упорядочена, т. е. строки ее переставлены в порядке возрастания



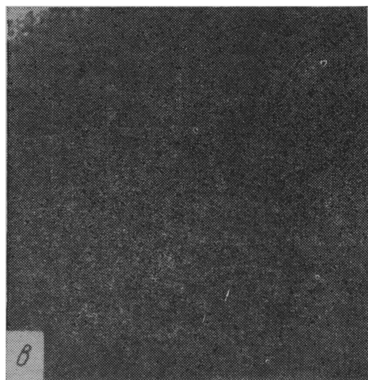


Рис. 10.5.4. Преобразование Адамара изображения «Портрет».

а — исходное изображение; *б* — спектр Адамара в логарифмическом масштабе по оси амплитуд; *в* — спектр с ограниченными наибольшими гармониками

секвенты, то существует другая форма записи преобразования Адамара. В этом случае

$$\mathscr{F}(u,\ v)=\frac{1}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}F(j,\ k)(-1)^{q(j,\ k,\ u,\ v)}, \qquad (10.5.6)$$

где

$$q(j, k, u, v) = \sum_{i=0}^{n-1} [g_i(u) j_i + g_i(v) k_i], \qquad (10.5.7)$$

причем

$$\begin{aligned} g_0(u) &\equiv u_{n-1}, \\ g_1(u) &\equiv u_{n-1} + u_{n-2}, \\ g_2(u) &\equiv u_{n-2} + u_{n-3}, \\ &\ldots\ldots\ldots\ldots \\ g_{n-1}(u) &\equiv u_1 + u_0. \end{aligned} \qquad (10.5.8)$$

Графики базисных функций преобразования Адамара с упорядоченной матрицей для $N = 16$ представлены на рис. 10.5.2. Базисные изображения, образованные с помощью матричного произведения базисных векторов преобразования Адамара размера 8×1, приведены на рис. 10.5.3. На рис. 10.5.4 дан пример преобразования Адамара (для упорядоченной матрицы Адамара).

10.6. ПРЕОБРАЗОВАНИЕ ХААРА

Преобразование Хаара [1, 23] основывается на ортогональной матрице Хаара. Ниже приведены примеры ортонормальных матриц Хаара четвертого и восьмого порядка:

$$\mathbf{H}_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}, \qquad (10.6.1)$$

$$\mathbf{H}_8 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}.$$

Матрицы Хаара более высокого порядка строятся по тем же правилам, что и матрицы (10.6.1) и (10.6.2). На рис. 10.6.1 приве-

дены графики базисных функций преобразования Хаара при $N = 16$. Базисные изображения того же преобразования для матрицы отсчетов, содержащей 8×8 элементов, представлены на рис. 10.6.2.

Преобразование Хаара можно рассматривать как процесс дискретизации исходного сигнала, при котором с переходом к следующей строке вдвое уменьшается шаг дискретизации.

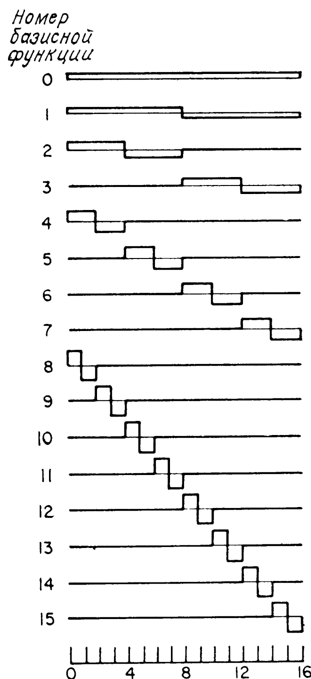


Рис. 10.6.1. Базисные функции преобразования Хаара при $N = 16$.

В задачах обработки изображений хааровский спектр описывает распределение энергии компонент, соответствующих разностям яркостей соседних элементов, разностям средних значений яркостей соседних пар элементов и вообще разностям средних значений яркостей соседних групп из 2^m элементов.

На рис. 10.6.3 приведен пример преобразования Хаара для конкретного изображения. На снимке с логарифмическим масштабом отчетливо заметна концентрация энергии, особенно в областях с высокими секвентами. Отметим, что в спектре Хаара наблюдается концентрация энергии также в областях с низкими секвентами.

10.7. НАКЛОННОЕ ПРЕОБРАЗОВАНИЕ

В работах [24—26] предложено ортогональное преобразование, названное наклонным (slant transform). Это преобразование обладает следующими особенностями: 1) среди базисных векторов

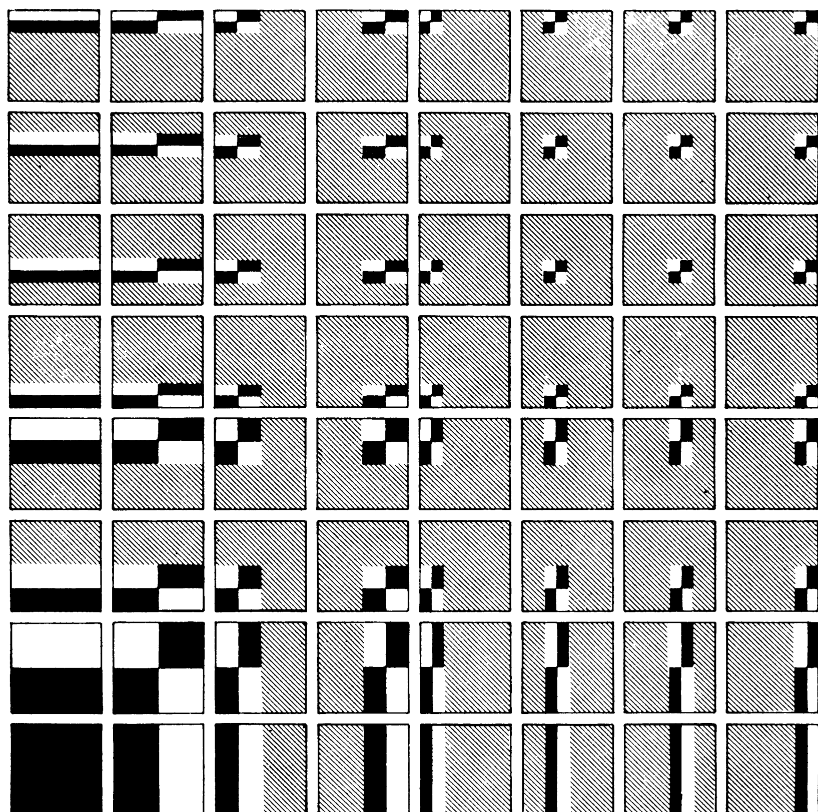
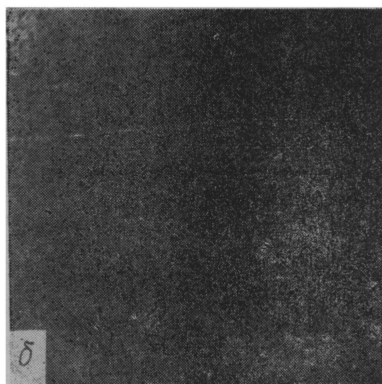


Рис. 10.6.2. Базисные изображения преобразования Хаара при $N = 8$. Черный цвет соответствует значению +1, белый — значению —1, штриховка — 0.

имеется вектор с одинаковыми компонентами (постоянный базисный вектор); 2) наклонный базисный вектор монотонно убывает от максимального до минимального значения скачками постоянной величины; 3) матрица преобразования обладает секвентным свойством; 4) существует быстрый алгоритм преобразования; 5) обеспечивается высокая степень концентрации энергии изображения. При длине вектора $N = 2$ наклонное преобразование совпадает

с преобразованием Адамара второго порядка. Таким образом,

$$\mathbf{S}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{10.7.1}$$



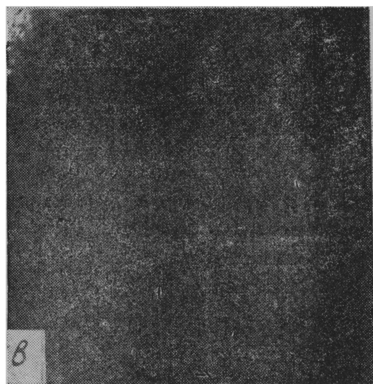


Рис. 10.6.3. Преобразование Хаара изображения «Портрет».

а — исходное изображение; *б* — спектр Хаара в логарифмическом масштабе по оси амплитуд; *в* — спектр с ограниченными наибольшими гармониками.

Матрица наклонного преобразования четвертого порядка формируется по следующему правилу:

$$\mathbf{S}_4 = \frac{1}{\sqrt{2}} \left[\begin{array}{cc:cc} 1 & 0 & 1 & 0 \\ a_4 & b_4 & -a_4 & b_4 \\ \hdashline 0 & 1 & 0 & -1 \\ -b_4 & a_4 & b_4 & a_4 \end{array} \right] \left[\begin{array}{c:c} \mathbf{S}_2 & 0 \\ \hdashline 0 & \mathbf{S}_2 \end{array} \right], \tag{10.7.2a}$$

или

$$\mathbf{S}_4=\frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & 1 & 1 \\ a_4+b_4 & a_4-b_4 & -a_4+b_4 & -a_4-b_4 \\ 1 & -1 & -1 & 1 \\ a_4-b_4 & -a_4-b_4 & a_4+b_4 & -a_4+b_4 \end{bmatrix}, \qquad (10.7.2б)$$

где a_4 и b_4 — действительные коэффициенты, которые следует выбирать так, чтобы матрица $\mathbf{S}_4$ была ортогональной, а величина скачков при изменении второго наклонного базисного вектора — постоянной. Из условия постоянства величины скачка можно найти, что $a_4 = 2b_4$. Из условия ортогональности $\mathbf{S}_4\mathbf{S}_4^T = \mathbf{I}$ следует, что $b_4 = 1/\sqrt{5}$. Таким образом, матрица наклонного преобразования четвертого порядка имеет вид

$$\mathbf{S}_4=\frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 3/\sqrt{5} & 1/\sqrt{5} & -1/\sqrt{5} & -3/\sqrt{5} \\ 1 & -1 & -1 & 1 \\ 1/\sqrt{5} & -3/\sqrt{5} & 3/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}. \qquad (10.7.3)$$

Нетрудно проверить, что матрица $\mathbf{S}_4$ является ортонормальной. Кроме того, она обладает секвентным свойством: число изменений знака возрастает с увеличением номера строки от 0 до 3.

Матрица наклонного преобразования при $N = 8$ имеет вид

$$\mathbf{S}_8=\frac{1}{2^{1/2}}\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ a_8 & b_8 & 0 & 0 & -a_8 & b_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ -b_8 & a_8 & 0 & 0 & b_8 & a_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{array}\right]\left[\begin{array}{c|c} \mathbf{S}_4 & 0 \\ \hline 0 & \mathbf{S}_4 \end{array}\right]. \qquad (10.7.4)$$

Как и при построении матрицы $\mathbf{S}_4$, коэффициенты a_8 и b_8 подбираются так, чтобы наклонный базисный вектор убывал равномерными скачками, все строки матрицы являлись ортонормальными векторами, а сама матрица обладала секвентным свойством.

Обобщая соотношение (10.7.4), можно получить рекуррентную формулу, связывающую матрицы наклонного преобразования

N-го и $(N/2)$-го порядка:

$$\mathbf{S}_N = \frac{1}{2^{1/2}} \left[\begin{array}{cc|c|cc|c} 1 & 0 & & 1 & 0 & \\ a_N & b_N & \mathbf{0} & -a_N & b_N & \mathbf{0} \\ \hline & \mathbf{0} & \mathbf{I}_{(N/2)-2} & & \mathbf{0} & \mathbf{I}_{(N/2)-2} \\ \hline 0 & 1 & & 0 & -1 & \\ -b_N & a_N & \mathbf{0} & b_N & a_N & \mathbf{0} \\ \hline & \mathbf{0} & \mathbf{I}_{(N/2)-2} & & \mathbf{0} & -\mathbf{I}_{(N/2)-2} \end{array}\right] \left[\begin{array}{c|c} \mathbf{S}_{N/2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{S}_{N/2} \end{array}\right], \quad (10.7.5)$$

где $\mathbf{I}_K$ — единичная матрица K-го порядка. Постоянные a_N и b_N можно найти из рекуррентных соотношений [26]

$$a_2 = 1, \quad (10.7.6а)$$

$$b_N = [1 + 4(a_{N/2})^2]^{-1/2}, \quad (10.7.6б)$$

$$a_N = 2b_N a_{N/2} \quad (10.7.6в)$$

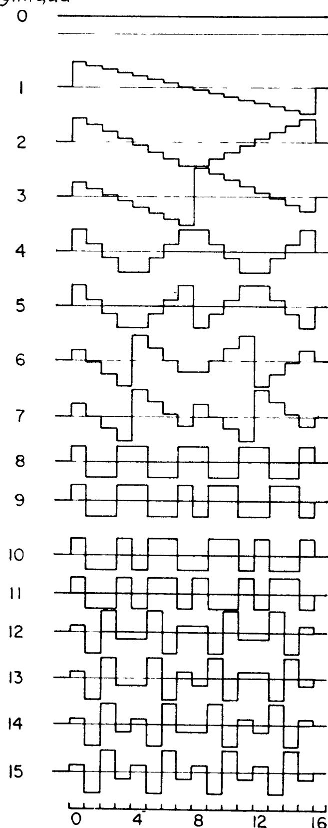


Рис. 10.7.1. Базисные функции наклонного преобразования при $N = 16$.

или по формулам

$$a_{2N} = \left(\frac{3N^2}{4N^2 - 1} \right)^{1/2}, \qquad (10.7.7а)$$

$$b_{2N} = \left(\frac{N^2 - 1}{4N^2 - 1} \right)^{1/2}. \qquad (10.7.7б)$$

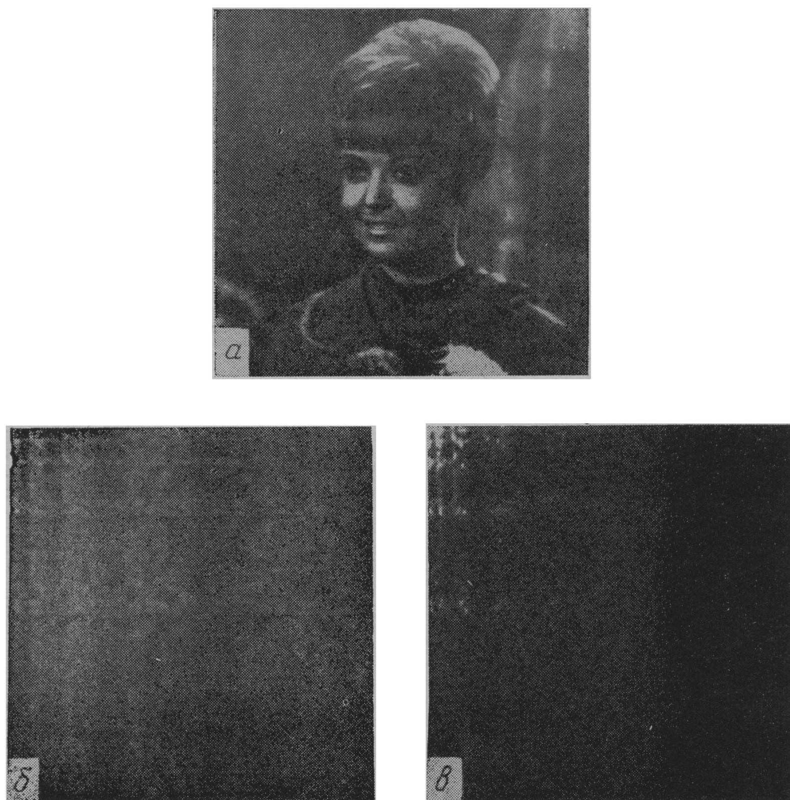


Рис. 10.7.2. Наклонное преобразование изображения «Портрет».

а — исходное изображение: *б* — результат преобразования в логарифмическом масштабе по оси амплитуд; *в* — результат преобразования с ограниченными наибольшими гармониками.

На рис. 10.7.1 приведены графики базисных функций наклонного преобразования для $N = 16$. Пример спектра, получаемого при таком преобразовании, показан на рис. 10.7.2.

10.8. ПРЕОБРАЗОВАНИЕ КАРУНЕНА—ЛОЭВА

Метод преобразования непрерывных сигналов в набор некоррелированных коэффициентов разработан Каруненом [27] и Лоэвом [28]. Как указывается в статье [30], Хотеллинг [29] первым предложил метод преобразования дискретных сигналов в набор некоррелированных коэффициентов. Однако в большинстве работ по цифровой обработке сигналов и дискретное, и непрерывное преобразования называют преобразованием Карунена—Лоэва или разложением по собственным векторам.

В общем случае преобразование Карунена—Лоэва описывается соотношением

$$\mathcal{F}(u, v) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) A(j, k; u, v), \tag{10.8.1}$$

ядро $A(j, k; u, v)$ которого удовлетворяет уравнению

$$\lambda(u, v) A(j, k; u, v) = \sum_{j'=0}^{N-1} \sum_{k'=0}^{N-1} K_F(j, k; j', k') A(j', k'; u, v), \tag{10.8.2}$$

где $K_F(j, k; j', k')$ — ковариационная функция дискретизованного изображения, а $\lambda(u, v)$ при фиксированных u и v постоянна. Функции $A(j, k; u, v)$ являются собственными функциями ковариационной функции, а $\lambda(u, v)$ — ее собственные значения. Как правило, выразить собственные функции в явной форме не удается.

Если ковариационную функцию можно разделить, т. е.

$$K_F(j, k; j', k') = K_C(j, j') K_R(k, k'), \tag{10.8.3}$$

то ядро разложения Карунена—Лоэва также разделимо и

$$A(j, k; u, v) = A_C(j, u) A_R(k, v). \tag{10.8.4}$$

Строки и столбцы матриц, описывающих эти ядра, удовлетворяют следующим уравнениям:

$$\lambda_R(v) A_R(k, v) = \sum_{k'=0}^{N-1} K_R(k, k') A_R(k', v), \tag{10.8.5}$$

$$\lambda_C(u) A_C(j, u) = \sum_{j'=0}^{N-1} K_C(j, j') A_C(j', u). \tag{10.8.6}$$

В частном случае, когда ковариационная матрица описывает разделимый марковский процесс первого порядка, собственные функции удается записать в явной форме. Для одномерного марковского процесса с коэффициентом корреляции ρ собственные функции и собственные значения имеют вид [3]

$$A(j, u) = \left[\frac{2}{N + \lambda^2(u)}\right]^{1/2} \sin\left[w(u)\left(j - \frac{N-1}{2}\right) + \frac{(u+1)\pi}{2}\right] \tag{10.8.7}$$

и

$$\lambda(u) = \frac{1 - \rho^2}{1 - 2\rho\cos[w(u)] + \rho^2}, \qquad (10.8.8)$$

где $0 \leqslant j,\ u \leqslant N - 1$, а $w(u)$ — корни трансцендентного уравнения

$$\operatorname{tg}(Nw) = \frac{(1 - \rho^2)\sin(w)}{\cos(w) - 2\rho + \rho^2\cos(w)}. \qquad (10.8.9)$$

Собственные векторы можно также найти из рекуррентных формул [32]

$$A(0, u) = \frac{\lambda(u)}{1 - \rho^2}[A(0, u) - \rho A(1, u)], \qquad (10.8.10а)$$

$$A(j, u) = \frac{\lambda(u)}{1 - \rho^2}[-\rho A(j-1, u) + (1 + \rho^2)A(j, u) - \rho A(j+1, u)],$$
$$0 < j < N - 1, \qquad (10.8.10б)$$

$$A(N-1, u) = \frac{\lambda(u)}{1 - \rho^2}[-\rho A(N-2, u) + A(N-1, u)], \qquad (10.8.10в)$$

положив в качестве начального условия $A(0, u) = 1$ и затем пронормировав полученные собственные векторы.

Если исходное и преобразованное изображения представить в векторной форме, то пара преобразований Карунена—Лоэва будет иметь вид

$$\mathscr{f} = \mathbf{A}\mathbf{f} \qquad (10.8.11)$$

и

$$\mathbf{f} = \mathbf{A}^T\mathscr{f}. \qquad (10.8.12)$$

Матрица преобразования $\mathbf{A}$ удовлетворяет уравнению

$$\mathbf{A}\mathbf{K}_f = \mathbf{\Lambda}\mathbf{A}, \qquad (10.8.13)$$

где $\mathbf{K}_f$ — ковариационная матрица вектора $\mathbf{f}$; $\mathbf{A}$ — матрица, строки которой являются собственными векторами матрицы $\mathbf{K}_f$; $\mathbf{\Lambda}$ — диагональная матрица вида

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda(1) & 0 & \cdots & 0 \\ 0 & \lambda(2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda(N^2) \end{bmatrix}. \qquad (10.8.14)$$

Если матрица $\mathbf{K}_f$ разделима, то

$$\mathbf{A} = \mathbf{A}_C \otimes \mathbf{A}_R, \qquad (10.8.15)$$

причем матрицы $\mathbf{A}_C$ и $\mathbf{A}_R$ удовлетворяют следующим условиям:

$$\mathbf{A}_R \mathbf{K}_R = \Lambda_R \mathbf{A}_R, \tag{10.8.16а}$$

$$\mathbf{A}_C \mathbf{K}_C = \Lambda_C \mathbf{A}_C, \tag{10.8.16б}$$

а $\lambda(k) = \lambda_R(i)\,\lambda_C(j)$ при $i, j = 1, 2, \ldots, N$ [33].

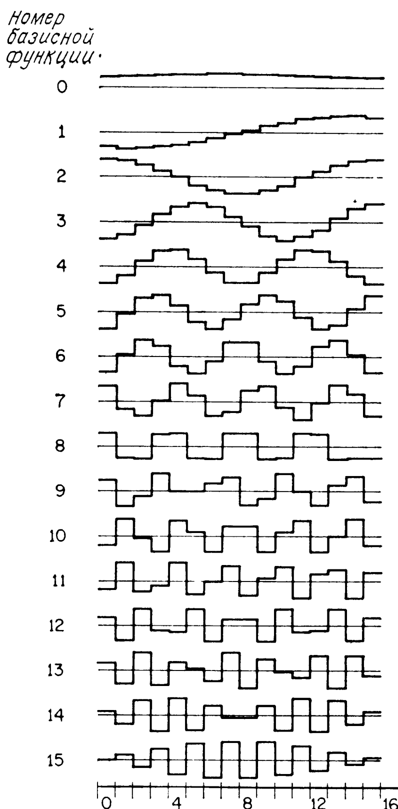
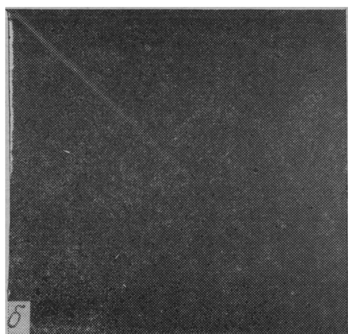


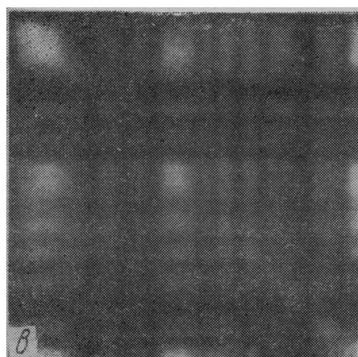
Рис. 10.8.1. Базисные функции преобразования Карунена — Лоэва при $N = 16$.

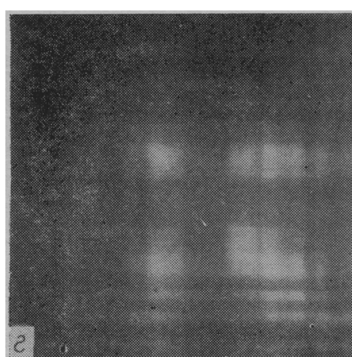
На рис. 10.8.1 приведены графики базисных функций преобразования Карунена—Лоэва одномерного марковского процесса, для которого коэффициенты корреляции соседних элементов $\rho = 0{,}9$.

10.9. СИНГУЛЯРНОЕ ПРЕОБРАЗОВАНИЕ

Сингулярное преобразование [34] является двумерным унитарным преобразованием, основанным на сингулярном разложе







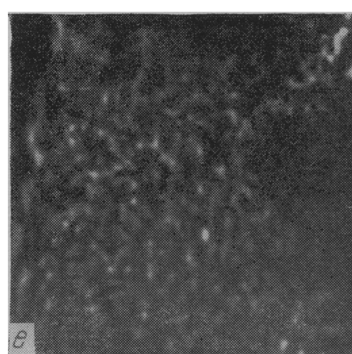


Рис. 10.9.1. Сингулярное преобразование изображения «Портрет». Все представленные массивы из 256×256 элементов получены из массивов размера 64×64 элемента посредством билинейной интерполяции.

а — исходное изображение, матрица $\mathbf{F}$; *б* — результат сингулярного преобразования, матрица $\mathcal{F}$; *в* — матрица $\mathbf{F}\mathbf{F}^T$; *г* — матрица $\mathbf{F}^T\mathbf{F}$; *д* — матрица, состоящая из модулей элементов матрицы $\mathbf{U}$; *е* — матрица, состоящая из модулей элементов матрицы $\mathbf{V}$.

нии матриц (см. гл. 5). Прямое сингулярное преобразование по определению равно

$$\mathcal{F} = \mathbf{A}_C \mathbf{F} \mathbf{A}_R^T, \tag{10.9.1}$$

а обратное преобразование

$$\mathbf{F} = \mathbf{A}_C^T \mathcal{F} \mathbf{A}_R. \tag{10.9.2}$$

Матрица преобразования строк $\mathbf{A}_R$ обеспечивает выполнение операции диагонализации

$$\mathbf{A}_R [\mathbf{F}^T \mathbf{F}] \mathbf{A}_R^T = \mathbf{\Lambda}, \tag{10.9.3}$$

где $\mathbf{\Lambda}$ — диагональная матрица, элементы $\lambda(u)$ которой являются собственными значениями матрицы $\mathbf{F}^T\mathbf{F}$. Аналогично

$$\mathbf{A}_C [\mathbf{F}\mathbf{F}^T] \mathbf{A}_C^T = \mathbf{\Lambda}. \tag{10.9.4}$$

Подставив выражение (10.9.2) в равенства (10.9.3) и (10.9.4), получим

$$\mathcal{F} = \mathbf{\Lambda}^{1/2}, \tag{10.9.5}$$

где элементами диагональной матрицы $\mathbf{\Lambda}^{1/2}$ являются числа $\lambda^{1/2}(u)$, называемые сингулярными значениями матрицы $\mathbf{F}$ и равные квадратному корню из соответствующих собственных значений $\lambda(u)$.

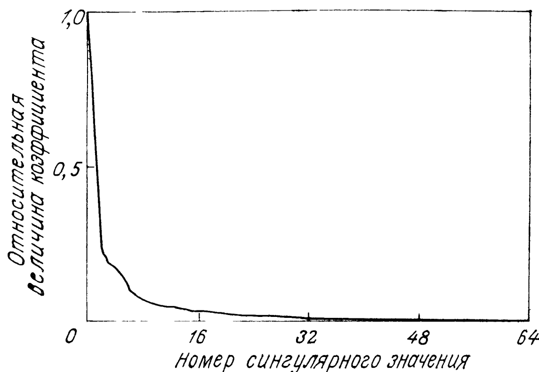
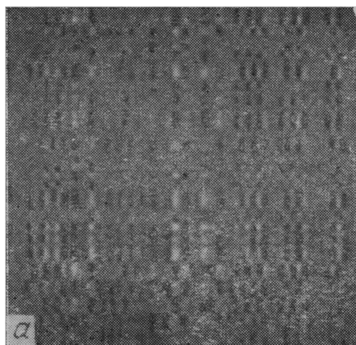


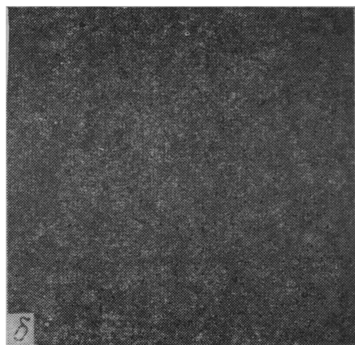
Рис. 10.9.2. Сингулярные значения изображения «Портрет».

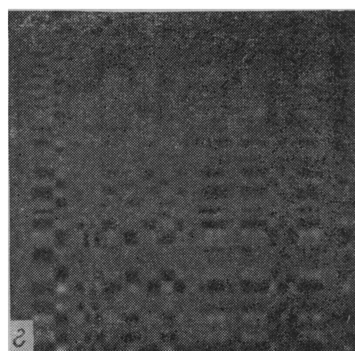
Матрицу изображения можно записать в очень компактной форме с помощью матричного произведения векторов, получаемых при сингулярном разложении. Согласно равенству (10.1.14б),

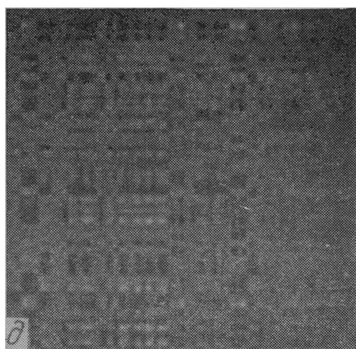
$$\mathbf{F} = \sum_{u=1}^{N} \mathcal{F}(u, u)\, \mathbf{b}_C(u)\, \mathbf{b}_R^T(u), \tag{10.9.6}$$

где $\mathbf{b}_C(u)$ и $\mathbf{b}_R(u)$ представляют собой векторы, состоящие из элементов u-х столбцов матриц $\mathbf{B}_C^T$ и $\mathbf{B}_R^T$.









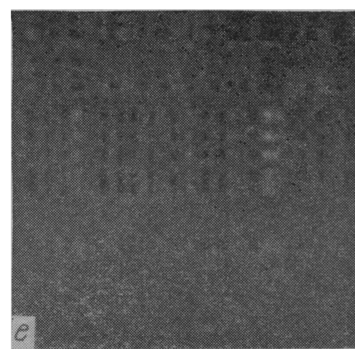


Рис. 10.9.3. Базисные изображения для изображения «Портрет».

Рисунки *а*, *б*, *в*, *г*, *д*, *е* соответствуют $b_C(1)\, b_R^T(1)$, $b_C(4)\, b_R^T(4)$, $b_C(8)\, b_R^T(8)$, $b_C(16)\, b_R^T(16)$, $b_C(32)\, b_R^T(32)$, $b_C(64)\, b_R^T(64)$.

Пользуясь сингулярным преобразованием, матрицу изображения $\mathbf{F}$, содержащую $N \times N$ элементов, можно полностью описать N величинами, представляющими собой N коэффициентов $\mathcal{F}(u, u)$. Однако следует отметить, что конкретные значения элементов матриц преобразований по строкам и столбцам в этом случае зависят от элементов изображения.

На рис. 10.9.1 приведен пример сингулярного преобразования изображения. Здесь же показаны произведения $\mathbf{F}\mathbf{F}^T$ и $\mathbf{F}^T\mathbf{F}$, а также соответствующие матрицы преобразований по строкам ($\mathbf{U}$) и по столбцам ($\mathbf{V}$). Сингулярные значения рассматриваемого изображения представлены на рис. 10.9.2. На рис. 10.9.3 показано несколько матричных произведений $\mathbf{b}_C(u)\,\mathbf{b}_R^T(u)$.

10.10. АЛГОРИТМЫ ВЫЧИСЛЕНИЙ

В общем случае, чтобы выполнить унитарное преобразование матрицы изображения, содержащей $N \times N$ элементов, и получить матрицу из $N \times N$ спектральных коэффициентов, необходимо произвести примерно N^4 арифметических операций (умножений и сложений). Если размеры матрицы изображения велики, то число операций становится чрезмерно большим. К счастью, для многих унитарных преобразований существуют эффективные алгоритмы вычислений, позволяющие ускорить выполнение преобразования.

Основной идеей этих быстрых вычислительных алгоритмов является разделение всей задачи на ряд этапов, причем результаты, полученные на предыдущих этапах, многократно используются на последующих этапах. В качестве примера рассмотрим процесс вычисления коэффициентов преобразования Адамара с неупорядоченной матрицей для последовательности из четырех элементов $f(j)$. При прямом способе вычисления находятся четыре величины по формулам

$$\mathcal{f}(0) = f(0) + f(1) + f(2) + f(3), \qquad (10.10.1а)$$
$$\mathcal{f}(1) = f(0) - f(1) + f(2) - f(3), \qquad (10.10.1б)$$
$$\mathcal{f}(2) = f(0) + f(1) - f(2) - f(3), \qquad (10.10.1в)$$
$$\mathcal{f}(3) = f(0) - f(1) - f(2) + f(3). \qquad (10.10.1г)$$

Для этого необходимо выполнить $[N(N-1)] = 12$ арифметических операций (сложений и вычитаний). Однако коэффициенты преобразования Адамара можно найти по-другому, разбив процесс вычисления на следующие этапы:

Первый этап

$$a(0) = f(0) + f(2), \qquad (10.10.2а)$$
$$a(1) = f(0) - f(2), \qquad (10.10.2б)$$
$$a(2) = f(1) + f(3), \qquad (10.10.2в)$$
$$a(3) = f(1) - f(3). \qquad (10.10.2г)$$

Второй этап

$$\mathscr{f}(0) = a(0) + a(2), \qquad (10.10.3а)$$

$$\mathscr{f}(1) = a(0) - a(2), \qquad (10.10.3б)$$

$$\mathscr{f}(2) = a(1) + a(3), \qquad (10.10.3в)$$

$$\mathscr{f}(3) = a(1) - a(3). \qquad (10.10.3г)$$

При этом для определения элементов матрицы $\mathbf{H}_4$ требуется только $[N \log_2 N] = 8$ операций, т. е. экономится четыре операции.

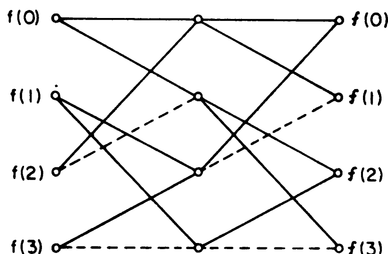


Рис. 10.10.1. Граф вычисления коэффициентов преобразования Адамара для четырехэлементной последовательности. (Сплошные линии обозначают операции сложения, пунктирные — вычитания.)

Ход вышеуказанных вычислений можно описать с помощью графа (рис. 10.10.1). Общее число операций равно половине числа ребер, соединяющих вершины графа. Другой способ представления того же процесса заключается в факторизации матрицы, когда матрицу преобразования Адамара $\mathbf{H}_4$ записывают в виде произведения двух разреженных матриц. Так, например,

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \qquad (10.10.4)$$

Число операций равно половине числа ненулевых элементов в обеих матрицах-сомножителях.

Принципы, описанные выше на примере преобразования Адамара, можно применить для быстрого вычисления многих других преобразований. Разработаны быстрые алгоритмы для преобразований Фурье [35], четного косинусного [12], синусного [13], Адамара [17], Хаара [1] и наклонного [26]. В общем случае для преобразования Карунена—Лоэва быстрого алгоритма не найдено, однако известны приближенные алгоритмы преобразования Карунена—Лоэва для марковских процессов.

Для большинства одномерных унитарных преобразований порядок требуемого числа арифметических операций равен $N \log_2 N$. Исключением является преобразование Хаара, которое имеет разреженную матрицу и поэтому может быть вычислено с помощью примерно $2N$ операций. Для унитарных преобразований пока не известен общий метод построения эффективных вычислительных алгоритмов [36]. В каждом случае приходится отыскивать эффективную процедуру вычислений или способ представления матрицы в виде произведения разреженных матриц.

Унитарные преобразования, основанные на базисных векторах синусоидального характера (преобразования Фурье, косинусное, синусное), можно выполнить косвенным путем, пользуясь алгоритмом так называемого z-преобразования с ЛЧМ-фильтрацией. Преоб-

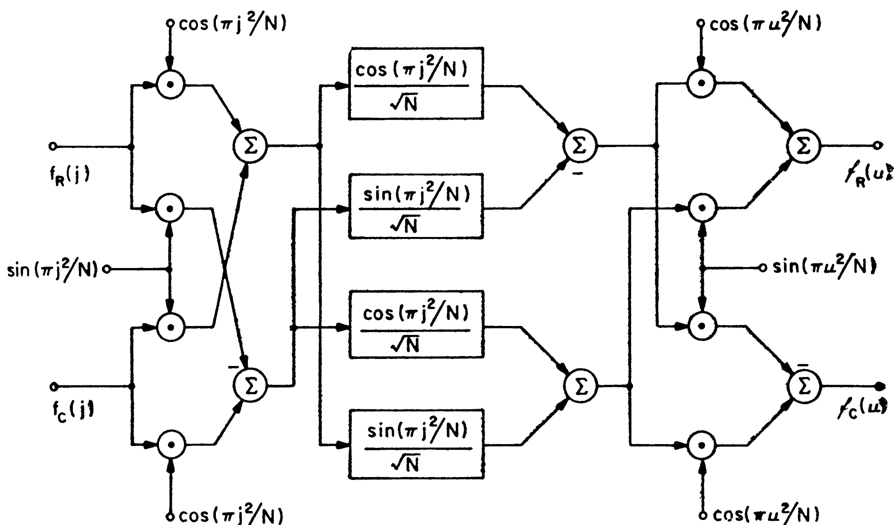


Рис. 10.10.2. Блок-схема алгоритма преобразования Фурье с использованием метода z-преобразования.

разование Фурье одномерной последовательности $f(j)$ определяется соотношением

$$\mathscr{f}(u) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f(j) \exp\left\{-\frac{2\pi i}{N} ju\right\}. \qquad (10.10.5)$$

Произведя замену переменных

$$2ju = j^2 + u^2 - (u - j)^2, \qquad (10.10.6)$$

можно получить выражение

$$\mathscr{f}(u) = \frac{\exp\{-i\pi u^2/N\}}{\sqrt{N}} \sum_{j=0}^{N-1} \exp\left\{\frac{i\pi}{N}(u-j)^2\right\} \exp\left\{-\frac{i\pi j^2}{N}\right\} f(j). \tag{10.10.7}$$

Формула (10.10.7) описывает совокупность следующих операций:

1. Поэлементное умножение последовательности $f(j)$ на множители с квадратично изменяющейся фазой $\exp\{-i\pi j^2/N\}$.

2. Свертка результатов предыдущей операции с ядром $\exp\{i\pi j^2/N\}$.

3. Поэлементное умножение полученной последовательности на множители с квадратичной фазой $\exp\{i\pi u^2/N\}$.

На рис. 10.10.2 приведена структурная схема алгоритма вычисления коэффициентов преобразования Фурье с использованием метода z-преобразования с ЛЧМ-фильтрацией. Если обработка производится в универсальной ЦВМ, то, как правило, применять этот алгоритм не следует, поскольку для вычисления свертки требуется большее число операций, чем для прямого выполнения преобразования методом БПФ. Однако с помощью аналоговых трансверсальных фильтров очень легко осуществить z-преобразование с ЛЧМ-фильтрацией, что дает возможность вычисления коэффициентов преобразования Фурье, а также косинусного и синусного преобразований поточным способом.

10.11. СТАТИСТИЧЕСКИЙ АНАЛИЗ ПРЕОБРАЗОВАННЫХ ИЗОБРАЖЕНИЙ

Для создания эффективных методов квантования и кодирования преобразованных изображений необходимо знать статистические свойства последних. В данном разделе определены первые и вторые моменты спектральных коэффициентов. Следующий раздел посвящен созданию статистической модели плотности вероятности этих коэффициентов.

Если предположить, что массив элементов изображения $F(j, k)$ является реализацией двумерного случайного процесса с известными математическим ожиданием и ковариационной функцией, то его спектр (результат унитарного преобразования)

$$\mathscr{F}(u, v) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) A(j, k; u, v) \tag{10.11.1}$$

также будет реализацией случайного процесса. Согласно равенству (8.2.1б), математическое ожидание $\mathscr{F}(u, v)$

$$E\{\mathscr{F}(u, v)\} = \sum_{j} \sum_{k} E\{F(j, k)\} A(j, k; u, v), \tag{10.11.2}$$

где $E\{F(j, k)\}$ — математическое ожидание $F(j, k)$. С помощью равенства (8.2.3) можно получить выражение для ковариационной функции спектра $\mathcal{F}(u, v)$

$$K_{\mathcal{F}}(u_1, v_1; u_2, v_2) = \sum_{j_1}\sum_{j_2}\sum_{k_1}\sum_{k_2} K_F(j_1, j_2; k_1, k_2) A(j_1, k_1; u_1, v_1) \times \times A^*(j_2, k_2; u_2, v_2), \quad (10.11.3)$$

где через $K_F(j_1, j_2; k_1, k_2)$ обозначена ковариационная функция $F(j, k)$. И наконец, дисперсия $\mathcal{F}(u, v)$ равна

$$\sigma^2_{\mathcal{F}}(u, v) = K_{\mathcal{F}}(u, v; u, v). \quad (10.11.4)$$

При векторном представлении унитарного преобразования вектор, образованный разверткой по столбцам матрицы коэффициентов преобразования, определяется соотношением

$$\mathscr{f} = \mathbf{A}\mathbf{f}. \quad (10.11.5)$$

Математическое ожидание этого вектора

$$\eta_{\mathscr{f}} = \mathbf{A}\eta_f, \quad (10.11.6)$$

а ковариационная матрица

$$\mathbf{K}_{\mathscr{f}} = \mathbf{A}\mathbf{K}_f\mathbf{A}^{*T}. \quad (10.11.7)$$

Матрица дисперсий коэффициентов преобразования $\mathbf{V}_{\mathcal{F}}$ образуется из диагональных элементов ковариационной матрицы (10.11.7).

Рассмотрим теперь, как вычисляются первый и второй моменты матрицы коэффициентов преобразования применительно к конкретным унитарным преобразованиям. Очевидно, что в общем случае все моменты можно получить в виде ряда или вектора, производя вычисления по вышеприведенным формулам. Однако для проектирования и анализа систем желательно иметь аналитические выражения моментных функций.

Сначала рассмотрим случай преобразования Карунена—Лоэва. В общем виде ковариационная функция коэффициентов этого преобразования описывается равенством (10.11.3), которое можно представить в следующем виде:

$$K_{\mathcal{F}}(u_1, v_1; u_2, v_2) = \sum_{j_2}\sum_{k_2} A^*(j_2, k_2; u_2, v_2) \sum_{j_1}\sum_{k_1} K_F(j_1, j_2; k_1, k_2) A(j_1, k_1; u_1, v_1). \quad (10.11.8)$$

Однако формула (10.8.2) подсказывает, что вторая двойная сумма совпадает с определением ядра преобразования Карунена—Лоэва. Таким образом,

$$K_{\mathcal{F}}(u_1, v_1; u_2, v_2) = \sum_{j_2}\sum_{k_2} A^*(j_2, k_2; u_2, v_2)\lambda(u_1, v_1) A(j_2, k_2; u_1, v_1), \quad (10.11.9)$$

где $\lambda(u, v)$ — собственные значения ковариационной матрицы изображения. Поскольку преобразование Карунена—Лоэва ортогонально, т. е.

$$K_{\mathscr{F}}(u_1, v_1;\ u_2, v_2) = \lambda(u_1, v_1)\,\delta(u_1 - u_2, v_1 - v_2), \quad (10.11.10)$$

то коэффициенты этого преобразования некоррелированы между собой, а их дисперсии равны соответствующим собственным значениям:

$$\sigma^2_{\mathscr{F}}(u, v) = \lambda(u, v). \quad (10.11.11)$$

Преобразование Карунена—Лоэва является единственным унитарным преобразованием, в котором достигается полная декорреляция произвольного изображения. В других преобразованиях между коэффициентами преобразования остается некоторая остаточная корреляция. Кроме того, преобразование Карунена—Лоэва обеспечивает наибольшую среди всех унитарных преобразований степень концентрации энергии спектра изображения. Предположим, что коэффициенты произвольного преобразования расположены в порядке убывания их дисперсий, т. е. $\sigma^2(1) > \sigma^2(2) > \ldots > \sigma^2(N^2)$, и также расположены собственные значения корреляционной матрицы, т. е. $\lambda(1) > \lambda(2) > \ldots > \lambda(N^2)$. Можно показать, что для любого верхнего предела суммирования $W < N^2$ выполняется неравенство

$$\sum_{w=0}^{W} \lambda(w) \geqslant \sum_{w=0}^{W} \sigma^2(w). \quad (10.11.12)$$

Чтобы найти аналитические выражения для моментов других преобразований, необходимо задать статистические свойства элементов изображения. Если изображение стационарно в широком смысле, то его математическое ожидание постоянно, и поэтому

$$E\{\mathscr{F}(u, v)\} = E\{\mathscr{F}\} \sum_j \sum_k A(j, k;\ u, v). \quad (10.11.13)$$

Но поскольку базисные функции являются ортогональными, то результат суммирования отличен от нуля только для нулевой базисной функции упорядоченного преобразования. Следовательно,

$$E\{\mathscr{F}(0, 0)\} = E\{F\} \sum_j \sum_k A(j, k;\ 0, 0) \text{ при } u, v = 0, \quad (10.11.14а)$$

$$E\{\mathscr{F}(u, v)\} = 0 \text{ при } u, v \neq 0. \quad (10.11.14б)$$

В тех преобразованиях, где в число базисных функций входит постоянная функция, двойная сумма (10.11.14а) равна N.

Если изображение стационарно в широком смысле, соответствующая ковариационная функция имеет вид $K_F(j_1 - j_2, k_1 - k_2)$. Однако матрица коэффициентов преобразования будет стационар-

ной в широком смысле только тогда, когда ядро преобразования обладает свойством пространственной инвариантности. Наглядным подтверждением этого является преобразование Фурье.

Большинство обычно используемых преобразований изображений являются разделимыми в пространственной области. Если, кроме того, разделима по пространственным переменным и ковариационная функция изображения, то

$$K_{\mathcal{F}}(u_1, v_1; u_2, v_2) = K_{\mathcal{F}_C}(u_1, u_2) K_{\mathcal{F}_R}(v_1, v_2), \quad (10.11.15)$$

где

$$K_{\mathcal{F}_C}(u_1, u_2) = \sum_{j_1} \sum_{j_2} K_{F_C}(j_1, j_2) A_C(j_1, u_1) A_C^*(j_2, u_2), \quad (10.11.16а)$$

$$K_{\mathcal{F}_R}(v_1, v_2) = \sum_{k_1} \sum_{k_2} K_{F_R}(k_1, k_2) A_R(k_1, v_1) A_R^*(k_2, v_2). \quad (10.11.16б)$$

При этом дисперсия спектра (преобразованного изображения) также разделима, т. е. ее можно представить в виде произведения одномерных функций, описывающих изменение дисперсии вдоль строк и столбцов:

$$\sigma^2_{\mathcal{F}}(u, v) = \sigma^2_{\mathcal{F}_R}(u)\, \sigma^2_{\mathcal{F}_C}(v). \quad (10.11.17)$$

Одномерная дисперсия Фурье-спектра для стационарного в широком смысле изображения имеет вид

$$\sigma^2_{\mathcal{F}_Z}(w) = \frac{1}{N} \sum_{z_1} \sum_{z_2} K_{F_Z}(z_1 - z_2) \exp\left\{-\frac{2\pi i}{N} w (z_1 - z_2)\right\}, \quad (10.11.18)$$

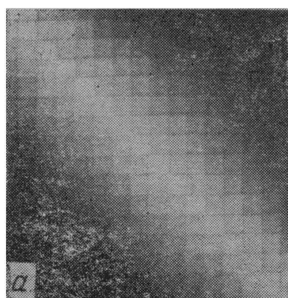
где $w = u$ или $w = v$, а $z_i = j_i$ или $z_i = k_i$. Это выражение можно переписать в следующем виде:

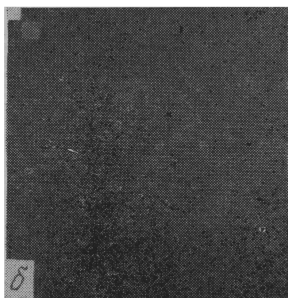
$$\sigma^2_{\mathcal{F}_Z}(w) = \frac{1}{N} \sum_{z_2} \exp\left\{\frac{2\pi i}{N} w z_2\right\} \sum_{z_1} K_{F_Z}\{z_1 - z_2\} \exp\left\{-\frac{2\pi i}{N} w z_1\right\}. \quad (10.11.19)$$

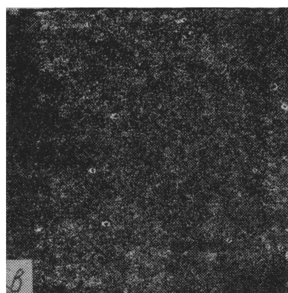
Вторую сумму можно рассматривать как одномерное дискретное преобразование Фурье ковариационной функции, сдвинутой на z_2 отсчетов. Согласно теореме о преобразовании Фурье смещенной функции,

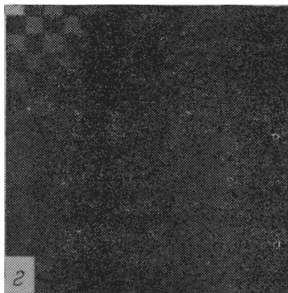
$$\sigma^2_{\mathcal{F}_Z}(w) = \frac{1}{N} \sum_{z_2} \exp\left\{\frac{2\pi i}{N} w z_2\right\} \exp\left\{-\frac{2\pi i}{N} w z_2\right\} \mathcal{G}_Z(w) = \mathcal{G}_Z(w), \quad (10.11.20)$$

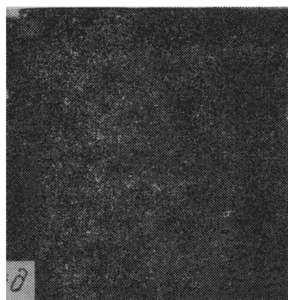
где $\mathcal{G}_Z(w)$ и $K_{F_Z}(w)$ — пара одномерных дискретных преобразо-

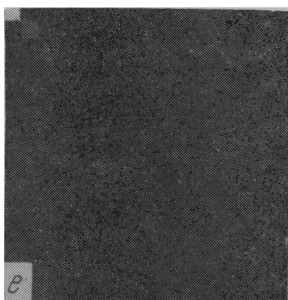


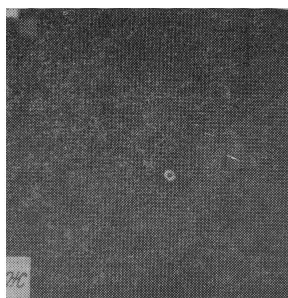












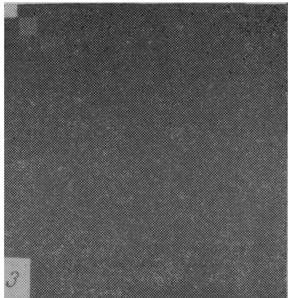


Рис. 10.11.1. Ковариационные матрицы унитарных преобразований при $N = 16$, $\rho = 0{,}9$.

а — тождественное преобразование; *б* — преобразование Фурье, *в* — косинусное; *г* — синусное; *д* — Адамара; *е* — Хаара; *ж* — наклонное; *з* — Карунена — Лоэва

ваний Фурье. Из соотношения (5.4.16) видно, что $\mathcal{G}_Z(w)$ является дискретным вариантом энергетического спектра $\mathcal{W}_{F_Z}(w)$ строки (или столбца) изображения, уменьшенным на величину, равную средней мощности изображения. Для процесса с нулевым математическим ожиданием

$$\sigma^2_{\mathcal{F}_Z}(w) = \mathcal{W}_{F_Z}(w). \qquad (10.11.21)$$

Рассмотренные ранее преобразования Адамара, Хаара и другие не обладают свойством, которое для преобразования Фурье дается теоремой о спектре сдвинутой функции. По этой причине для этих преобразований не удалось получить в конечной форме выражение для дисперсии или ковариационной функции.

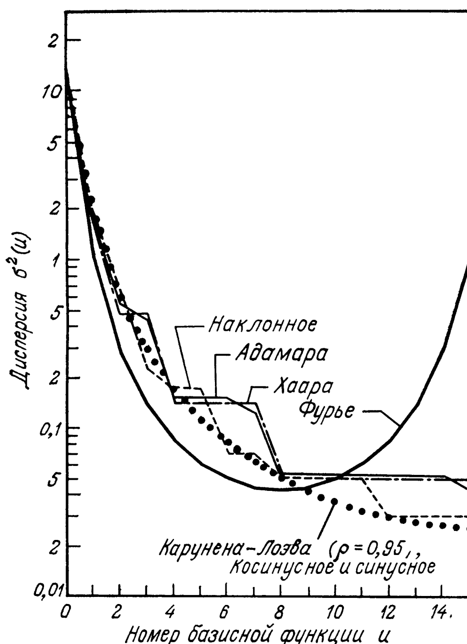


Рис. 10.11.2. Дисперсии коэффициентов унитарных преобразований при $N = 16$, $\rho = 0{,}9$.

На рис. 10.11.1 представлены ковариационные функции коэффициентов некоторых унитарных преобразований отрезка дискретного одномерного марковского процесса, состоящего из 16 элементов, причем коэффициент корреляции между соседними элементами $\rho = 0{,}9$. Графики дисперсии коэффициентов преобразований приведены на рис. 10.11.2.

10.12. МОДЕЛИ ПЛОТНОСТИ ВЕРОЯТНОСТИ ПРЕОБРАЗОВАННЫХ ИЗОБРАЖЕНИЙ

Было бы полезно знать плотность вероятности элементов преобразованного изображения для произвольного преобразования. К сожалению, найти эту плотность нелегко, так как обычно мы не имеем полного описания плотности вероятности исходного изображения. Кроме того, сами преобразования, как правило, являются сложными математическими операциями. Однако рассмотренные выше преобразования, применяемые для обработки изображений, сводятся к вычислению взвешенных сумм всех элементов исходного изображения. Поэтому, пользуясь центральной предельной теоремой теории вероятностей, можно прийти к выводу о том, что плотность вероятности элементов преобразованного изображения близка к гауссовой с теми моментами, которые были найдены в предыдущем разделе.

Коэффициенты преобразования Фурье являются комплексными числами, и их можно задавать с помощью действительной и мнимой частей или же модулем и аргументом. В обоих случаях каждому коэффициенту соответствуют два числа, которые необходимо подвергнуть квантованию. Можно считать, что действительная $\mathcal{F}_R(u, v)$ и мнимая $\mathcal{F}_I(u, v)$ части коэффициента Фурье имеют одинаковые гауссовы распределения, дисперсия которых $\sigma^2_{\mathcal{F}}(u, v)$ пропорциональна энергетическому спектру исходного изображения. Таким образом,

$$p\{\mathcal{F}_R(u, v)\} = [2\pi\sigma^2_{\mathcal{F}}(u, v)]^{-1/2} \exp\left\{-\frac{\mathcal{F}^2_R(u, v)}{2\sigma^2_{\mathcal{F}}(u, v)}\right\}, \quad (10.12.1а)$$

$$p\{\mathcal{F}_I(u, v)\} = [2\pi\sigma^2_{\mathcal{F}}(u, v)]^{-1/2} \exp\left\{-\frac{\mathcal{F}^2_I(u, v)}{2\sigma^2_{\mathcal{F}}(u, v)}\right\}. \quad (10.12.1б)$$

Если действительная и мнимая составляющие имеют гауссово распределение, то модуль коэффициента Фурье $\mathcal{F}_M(u, v)$ распределен по закону Рэлея. Таким образом,

$$p\{\mathcal{F}_M(u, v)\} = \frac{\mathcal{F}_M(u, v)}{\sigma^2_{\mathcal{F}}(u, v)} \exp\left\{-\frac{\mathcal{F}^2_M(u, v)}{2\sigma^2_{\mathcal{F}}(u, v)}\right\} \quad (10.12.2)$$

при $\mathcal{F}_M(u, v) \geqslant 0$, а фаза $\mathcal{F}_P(u, v)$ имеет равномерное распределение

$$p\{\mathcal{F}_P(u, v)\} = 1/2\pi, \quad (10.12.3)$$

где $-\pi \leqslant \mathcal{F}_P(u, v) \leqslant \pi$.

Можно считать, что действительные коэффициенты других преобразований имеют гауссовы распределения вида

$$p\{\mathcal{F}(u, v)\} = [2\pi\sigma^2_{\mathcal{F}}(u, v)]^{-1/2} \exp\left\{-\frac{\mathcal{F}^2(u, v)}{2\sigma^2_{\mathcal{F}}(u, v)}\right\}. \quad (10.12.4)$$

Плотность распределения вероятностей Рэлея (10.12.2) часто применяют в качестве модели плотности вероятности постоянной составляющей $\mathcal{F}(0, 0)$ яркостных изображений, поскольку для таких изображений коэффициент $\mathcal{F}(0, 0)$ всегда строго положителен. Рассмотренные модели плотностей вероятности широко применяются при разработке методов кодирования с преобразованием изображений.

ЛИТЕРАТУРА

1. Andrews H. C., Computer Techniques in Image Processing, Academic Press, New York, 1970.
2. Andrews H. C., Two Dimensional Transforms, in: Topics in Applied Physics: Picture Processings and Digital Filtering, Vol. 6, Huang T., Ed., Springer-Verlag, New York, 1975.
3. Bellman R., Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.
4. Andrews H. C., Caspari K., A Generalized Technique for Spectral Analysis, *IEEE Trans. Computers*, **C-19**, 1, 16—25 (January 1970).
5. Cooley J. W., Tukey J. W., An Algorithm for the Machine Calculation of Complex Fourier Series, *Math. Comput.*, **19**, 90, 297—301 (April 1965).
6. IEEE Trans. Audio and Electroacoustics, Special Issue on Fast Fourier Transforms, **AU-15**, 2 (June 1967).
7. Cochran W. T. et al., What is the Fast Fourier Transform? *Proc. IEEE*, **55**, 10, 1664—1674 (1967). [Имеется перевод: Кокрен и др., Что такое «Быстрое преобразование Фурье»? — ТИИЭР, 1967, т. 55, № 10.]
8. *IEEE Trans. Audio and Electroacoustics*, Special Issue on Fast Fourier Transforms, **AU-17**, 2 (June 1969).
9. Cooley J. W., Lewis P. A., Welch P. D., Historical Notes on the Fast Fourier Transform, *Proc. IEEE*, **55**, 10, 1675—1677 (October 1967). [Имеется перевод: Кули, Льюис, Уэлч, Исторические замечания относительно быстрого преобразования Фурье. — ТИИЭР, 1967, т. 55, № 10.]
10. Brigham E. O., Morrow R. E., The Fast Fourier Transform, *IEEE Spectrum*, **4**, 12, 63—70 (December 1967).
11. Means R. W., Whitehouse H. J., Speiser J. M., Television Encoding Using a Hybrid Discrete Cosine Transform and a Differential Pulse Code Modulator in Real Time, Proceedings of the National Telecommunications Conference, San Diego, California, December 1974, pp. 61—66.
12. Ahmed N., Natarajan T., Rao K. R., On Image Processing and a Discrete Cosine Transform, *IEEE Trans. Computers*, **C-23**, 1, 90—93 (January 1974).
13. Jain A. K., A Fast Karhunen-Loeve Transform for Finite Discrete Images, Proceedings of the National Electronics Conference, Chicago, Illinois, October 1974, pp. 323—328.
14. Jain A. K., Angel E., Image Restoration, Modelling, and Reduction of Dimensionality, *IEEE Trans. Computers*, **C-23**, 5, 470—476 (May 1974).
15. Jain A. K. (частное сообщение).
16. Whelchel J. E., Jr., Guinn D. F., The Fast Fourier-Hadamard Transform and Its use in Signal Representation and Classification, EASCON 1968 Convention Record, 1968, pp. 561—573.
17. Pratt W. K., Andrews H. C., Kane J., Hadamard Transform Image Coding, *Proc. IEEE*, **57**, 1, 58—68 (January 1969). [Имеется перевод: Прайт, Кайн, Эндрюс, Кодирование изображений с использованием преобразования Адамара. — ТИИЭР, 1969, т. 57, № 1.]
18. Hadamard J., Resolution d'une Question Relative aux Determinants, *Bull. Sci. Math. Ser.* 2, **17**, Part I, 240—246 (1893).

19. Ryser H. J., Combinational Mathematics, Wiley, New York, 1963. [Имеется перевод: Райзер Г. Дж., Комбинаторная математика. — М.: Мир, 1966.]
20. Harmuth H. F., Transmission of Information by Orthogonal Functions, Springer-Verlag, New York, 1969. [Имеется перевод: Хармут Х. Ф., Передача информации ортогональными функциями. — М.: Связь, 1975.]
21. Walsh J. L., A Closed Set of Orthogonal Functions, *Am. J. Math.*, **45**, 5—24 (1923).
22. Rademacher H., Einige Satze von Allgemeinen Orthogonal-Funktionen, *Math. Annalen*, **87**, 112—138 (1922).
23. Haar A., Zur Theorie der Orthogonalen Funktionen-System, Inaugural Dissertation, *Math. Annalen*, **5**, 17—31 (1955).
24. Enomoto H., Shibata K., Orthogonal Transform Coding System for Television Signals, *IEEE Trans. Electromagnetic Compatibility*, **EMC-13**, 3, 11—17 (August 1971).
25. Pratt W. K., Welch L. R., Chen W. H., Slant Transform for Image Coding, Proceedings Symposium an Applications of Walsh Functions, March 1972.
26. Pratt W. K., Chen W. H., Welch L. R., Slant Transform Image Coding, *IEEE Trans. Commun.*, **COM-22**, 8, 1075—1093 (August 1974).
27. Karhunen H., 1947, English translation by Selin I., On Linear Methods in Probability Theory, The Rand Corporation, Doc. T-131, August 1960.
28. Loeve M., Fonctions Aléatories de Seconde Ordre, Hermann, Paris, 1948.
29. Hotelling H., Analysis of a Complex of Statistical Variables into Principal Components, *J. Educ. Psych.*, **24**, 417—441, 498—520 (1933).
30. Wintz P. A., Transform Picture Coding, *Proc. IEEE*, **60**, 7, 809—820 (July 1972). [Имеется перевод: Уинтц, Кодирование изображения посредством преобразований. — ТИИЭР, 1972, т. 60, № 7.]
31. Ray W. D., Driver R. M., Further Decomposition of the Karhunen-Loeve Series Representation of a Stationary Random Process, *IEEE Trans. Inf. Theory*, **IT-16**, 6, 663—668 (November 1970).
32. Pratt W. K., Generalized Wiener Filtering Computation Techniques, *IEEE Trans. Computers*, **C-21**, 7, 636—641 (July 1972).
33. Friedman B., Eigenvalues of Composite Matrices, *Proc. Cambridge Phil. Soc.*, **57**, Part 1, 37—49 (1961).
34. Andrews H. C., Patterson C. L., Outer Product Expansions and Their Uses in Digital Image Processing, *Am. Math. Monthly*, **1**, 82, 1—13 (January 1975).
35. Brigham E. O., The Fast Fourier Transform, Prentice-Hall, Englewood Cliffs, N. J., 1974.
36. Gentleman W. M., Matrix Multiplication and Fast Fourier Transformations, *Bell. Syst. Tech. J.*, **47**, 1099—1103 (July — August 1968).

Глава 11

ДВУМЕРНЫЕ МЕТОДЫ ЛИНЕЙНОЙ ОБРАБОТКИ

Большинство алгоритмов обработки дискретных изображений являются линейными: элементы обработанного изображения образуются в виде линейной комбинации элементов исходного изображения. Широкое применение линейных операций объясняется тем, что выполнить их легче, чем нелинейные. Однако при обработке массивов большого размера часто не удается выполнить даже линейные операции, если отсутствуют эффективные алгоритмы вычисления. В данной главе рассмотрены косвенные методы вычислений, опирающиеся на унитарные преобразования и позволяющие проводить обработку более эффективно, чем при использовании традиционных методов. Здесь также описаны методы рекурсивной двумерной фильтрации.

11.1. ОБРАБОТКА С ИСПОЛЬЗОВАНИЕМ ПРЕОБРАЗОВАНИЯ

Двумерные линейные преобразования можно определить в виде ряда как

$$P(m_1, m_2) = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} F(n_1, n_2)\, T(n_1, n_2; m_1, m_2). \qquad (11.1.1)$$

Это соотношение в векторной форме имеет вид

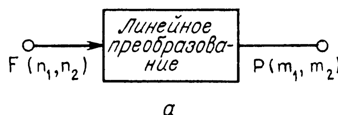
$$\mathbf{p} = \mathbf{T}\mathbf{f}. \qquad (11.1.2)$$

Ниже будет показано, что подобные линейные преобразования часто можно выполнить более эффективно, если вместо непосредственных вычислений по формулам (11.1.1) или (11.1.2) воспользоваться косвенными методами с применением двумерных унитарных преобразований.

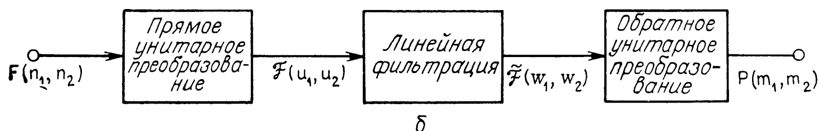
На рис. 11.1.1 приведена структурная схема косвенного метода вычислений, называемого обобщенной линейной фильтрацией [1]. Массив $F(n_1, n_2)$, описывающий исходное изображение, подвергается здесь двумерному унитарному преобразованию, в результате которого получается массив коэффициентов преобразования $\mathcal{F}(u_1, u_2)$. Затем составляется линейная комбинация этих коэффициентов, описываемая в общем случае формулой

$$\tilde{\mathcal{F}}(w_1, w_2) = \sum_{u_1=1}^{N_1} \sum_{u_2=1}^{N_2} \mathcal{F}(u_1, u_2)\, \mathcal{T}(u_1, u_2; w_1, w_2), \qquad (11.1.3)$$

где $\mathscr{T}(u_1, u_2; w_1, w_2)$ — ядро линейного фильтрующего преобразования. В заключение проводится обратное унитарное преобразование, чтобы получить обработанное изображение $P(m_1, m_2)$.



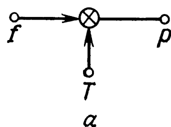
а



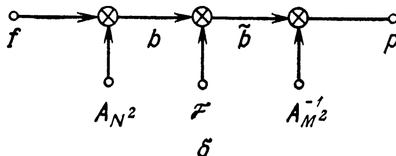
б

Рис. 11.1.1. Структура алгоритмов непосредственной обработки (*а*) и обобщенной линейной фильтрации (*б*).

Такой метод вычисления будет эффективнее, чем непосредственное вычисление по формуле (11.1.1), если существуют быстрые алгоритмы унитарного преобразования, а ядро $\mathscr{T}(u_1, u_2; w_1, w_2)$ содержит большое количество нулевых элементов.



а



б

Рис. 11.1.2. Структура алгоритмов непосредственной обработки (*а*) и обобщенной линейной фильтрации (*б*) при векторном представлении изображений.

Процесс обобщенной линейной фильтрации можно также представить в векторной форме (рис. 11.1.2). Для упрощения обозначений положим, что $N_1 = N_2 = N$, а $M_1 = M_2 = M$. Тогда обобщенную линейную фильтрацию можно описать соотношениями

$$\mathscr{f} = (\mathbf{A}_{N^2})\,\mathbf{f}, \quad (11.1.4а)$$

$$\tilde{\mathscr{f}} = \mathscr{T}\mathscr{f}, \quad (11.1.4б)$$

$$\mathbf{p} = (\mathbf{A}_{M^2})^{-1}\tilde{\mathscr{f}}, \quad (11.1.4в)$$

где $\mathbf{A}_{N^2}$ — матрица унитарного преобразования размера $N^2 \times N^2$, $\mathcal{T}$ — матрица оператора линейной фильтрации размера $M^2 \times N^2$, а $\mathbf{A}_{M^2}$ — матрица обратного унитарного преобразования размера $M^2 \times M^2$. Как следует из соотношений (11.1.4), векторы исходного и обработанного изображений связаны между собой равенством

$$\mathbf{p} = (\mathbf{A}_{M^2})^{-1} \mathcal{T} (\mathbf{A}_{N^2}) \mathbf{f}. \qquad (11.1.5)$$

Приравнивая соотношения (11.1.2) и (11.1.5), найдем связь между матрицами $\mathcal{T}$ и $\mathbf{T}$:

$$\mathbf{T} = (\mathbf{A}_{M^2})^{-1} \mathcal{T} (\mathbf{A}_{N^2}), \qquad (11.1.6а)$$

$$\mathcal{T} = (\mathbf{A}_{M^2}) \mathbf{T} (\mathbf{A}_{N^2})^{-1}. \qquad (11.1.6б)$$

При непосредственной обработке для выполнения вычислений по формуле (11.1.2) необходимо провести $k_d\ (M^2N^2)$ арифметических действий, где k_d $(0 \leqslant k_d \leqslant 1)$ характеризует степень заполнения матрицы $\mathbf{T}$ ненулевыми элементами. При обобщенной линейной фильтрации общее число арифметических действий равно:

Прямое преобразование: N^4 — при непосредственном вычислении; $2N^2 \log_2 N$ — при использовании быстрого преобразования.

Линейная фильтрация: $k_{\mathcal{T}} M^2N^2$ умножений.

Обратное преобразование: M^4 — при непосредственном вычислении; $2M^2 \log_2 M$ — при использовании быстрого преобразования.

Здесь $k_{\mathcal{T}}$ $(0 \leqslant k_{\mathcal{T}} \leqslant 1)$ — коэффициент заполнения матрицы $\mathcal{T}$ ненулевыми элементами. Если $k_{\mathcal{T}} = 1$ и производится непосредственное вычисление унитарного преобразования, то очевидно, что обобщенная линейная фильтрация оказывается не столь эффективной, как непосредственная обработка. Однако, если применяются быстрые алгоритмы, подобные быстрому преобразованию Фурье (БПФ), обобщенная линейная фильтрация будет эффективнее, чем непосредственная обработка, если коэффициент заполнения матрицы $\mathcal{T}$ удовлетворяет неравенству

$$k_g < k_d - (2/M^2) \log_2 N - (2/N^2) \log_2 M. \qquad (11.1.7)$$

Во многих случаях матрица $\mathcal{T}$ бывает разрежена и неравенство (11.1.7) выполняется. Вообще говоря, унитарные преобразования часто приводят к декорреляции элементов матрицы $\mathbf{T}$ и матрица $\mathcal{T}$ оказывается разреженной. Кроме того, матрицу $\mathcal{T}$ часто удается разредить без внесения больших погрешностей, если заменить нулями ее малые по величине элементы [1].

В последующих разделах рассмотрена структура линейных операторов суперпозиции, свертки и псевдообращения, для того чтобы определить возможность применения метода обобщенной линейной фильтрации для осуществления этих операций.

11.2. СУПЕРПОЗИЦИЯ С ПРЕОБРАЗОВАНИЕМ

Операцию суперпозиции, описанную в гл. 9, часто удается выполнить более эффективно, если вместо непосредственной обработки проводить обработку с использованием преобразования. На рис. 11.2.1, *а*, *б* приведены схемы непосредственного выполнения конечной суперпозиции и дискретизованной суперпозиции. На рис. 11.2.1, *г*, *д* представлены схемы осуществления операции суперпозиции, когда вектор $\mathbf{f}$ подвергается унитарному преобразованию, результат которого умножается на матрицу $\mathscr{D}$ (для оператора суперпозиции конечных массивов) или матрицу $\mathscr{B}$ (для дискретизованного оператора суперпозиции). Обратное преобразование дает обработанный вектор. Согласно рис. 11.2.1, *а*, *г*, в случае применения оператора суперпозиции конечных массивов исходный и обработанный векторы связаны следующими соотношениями:

$$\mathbf{q} = \mathbf{D}\mathbf{f} \tag{11.2.1а}$$

и

$$\mathbf{q} = (\mathbf{A}_{M^2})^{-1} \mathscr{D} (\mathbf{A}_{N^2}) \mathbf{f}. \tag{11.2.1б}$$

Поэтому матрицу $\mathscr{D}$ можно представить в виде произведения

$$\mathscr{D} = (\mathbf{A}_{M^2}) \mathbf{D} (\mathbf{A}_{N^2})^{-1}. \tag{11.2.2а}$$

Аналогично

$$\mathscr{B} = (\mathbf{A}_{M^2}) \mathbf{B} (\mathbf{A}_{N^2})^{-1}. \tag{11.2.2б}$$

Чтобы непосредственно выполнить суперпозицию конечных массивов, требуется примерно N^2L^2 операций (здесь L — порядок матрицы импульсного отклика). В этом случае коэффициент заполнения матрицы $\mathbf{D}$ равен

$$k_{\mathrm{D}} = (L/N)^2. \tag{11.2.3а}$$

Для суперпозиции дискретизованных массивов требуется около M^2L^2 операций, а коэффициент заполнения матрицы $\mathbf{B}$ равен

$$k_{\mathrm{B}} = (L/M)^2. \tag{11.2.3б}$$

На рис. 11.2.1, *е* приведена блок-схема циклической суперпозиции с преобразованием. В этом случае входным вектором $\mathbf{k}_E$ служит так называемый расширенный входной вектор, получаемый размещением матрицы исходного изображения $F(n_1, n_2)$ в левом углу квадратной нулевой матрицы J-го порядка и разверткой полученной матрицы по столбцам. Повторяя вышеприведенные рассуждения, можно показать, что

$$\mathbf{k}_E = \mathbf{C}\mathbf{f}_E = (\mathbf{A}_{J^2})^{-1} \mathscr{C} (\mathbf{A}_{J^2}) \mathbf{f}_E \tag{11.2.4а}$$

и, следовательно,

$$\mathscr{C} = (\mathbf{A}_{J^2}) \mathbf{C} (\mathbf{A}_{J^2})^{-1}. \tag{11.2.4б}$$

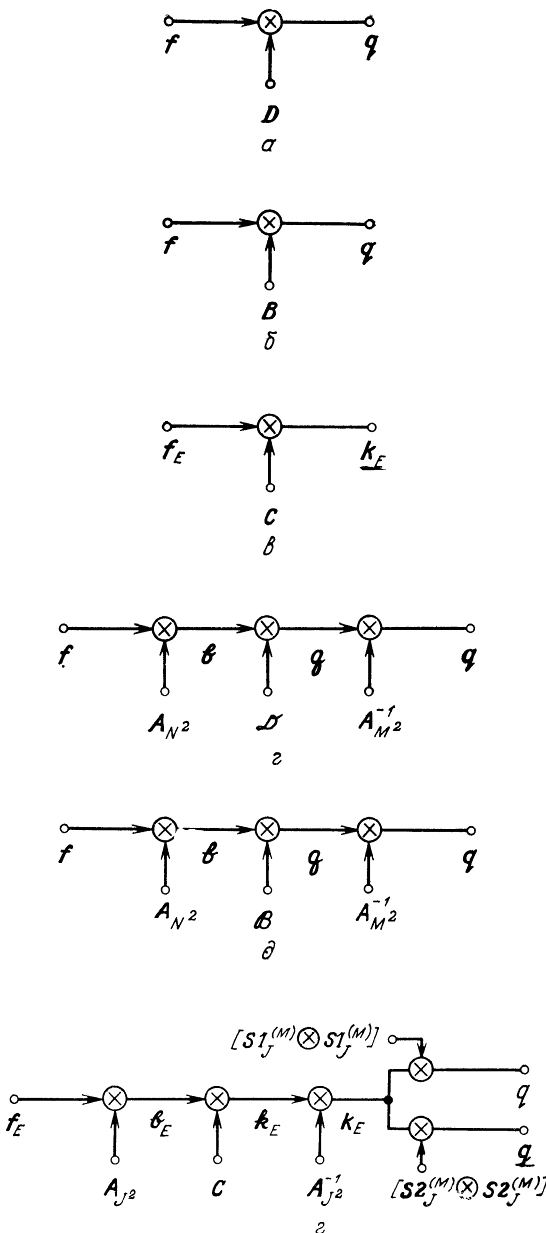


Рис. 11.2.1. Различные виды суперпозиции.

а — конечная суперпозиция, *б* — дискретизованная интегральная суперпозиция; *в* — циклическая суперпозиция; *г* — конечная суперпозиция с преобразованием; *д* — дискретизованная интегральная суперпозиция с преобразованием; *е* — циклическая суперпозиция с преобразованием.

Как было отмечено в гл. 9, для суперпозиции конечных или дискретизованных массивов эквивалентный выходной вектор можно сформировать из $\mathbf{k}_E$, выбирая определенные компоненты последнего. Для суперпозиции конечных массивов

$$\mathbf{q} = [\mathbf{S1}_J^{(M)} \otimes \mathbf{S1}_J^{(M)}]\,\mathbf{k}_E, \tag{11.2.5а}$$

а для суперпозиции дискретизованных массивов

$$\mathbf{q} = [\mathbf{S2}_J^{(M)} \otimes \mathbf{S2}_J^{(M)}]\,\mathbf{k}_E. \tag{11.2.5б}$$

При суперпозиции конечных массивов матрица обработанного изображения связана с расширенной матрицей изображения $\mathbf{K}_E$ следующим образом:

$$\mathbf{Q} = [\mathbf{S1}_J^{(M)}]\,\mathbf{K}_E\,[\mathbf{S1}_J^{(M)}]^T. \tag{11.2.6а}$$

При суперпозиции дискретизованных массивов матрица обработанного изображения равна

$$\mathbf{G} = [\mathbf{S2}_J^{(M)}]\,\mathbf{K}_E\,[\mathbf{S2}_J^{(M)}]^T. \tag{11.2.6б}$$

Число арифметических операций, необходимых для вычисления вектора $\mathbf{k}_E$ путем обработки с преобразованием, можно найти из приведенных выше формул, если положить $M = N = J$:

Прямое преобразование: $3J^4$.

Быстрое преобразование: $J^2 + 4J^2 \log_2 J$.

Если матрица $\mathscr{C}$ разрежена, то многие из J^2 умножений, нужных для операции фильтрации, выполняться не будут.

Из вышеизложенного нетрудно сделать вывод, что для эффективного проведения суперпозиции следует подобрать преобразование, отвечающее двум требованиям: во-первых, для него должен существовать быстрый алгоритм, а, во-вторых, матрица фильтрации преобразованных массивов должна быть разреженной. В качестве примера рассмотрим свертку конечных массивов, получаемую с помощью преобразования Фурье [2, 3]. В соответствии с рис. 11.2.1 положим

$$\mathbf{A}_{K^2} = \mathbf{A}_K \otimes \mathbf{A}_K, \tag{11.2.7}$$

где

$$\mathbf{A}_K = [(1/\sqrt{K})\,\mathscr{W}^{(x-1)\,(y-1)}], \quad \mathscr{W} \equiv \exp\left\{-\frac{2\pi i}{K}\right\}$$

при $x, y = 1, 2, \ldots, K$. Кроме того, примем, что вектор $\mathbf{h}_E^{(K)}$ размера $K^2 \times 1$ представляет расширенную матрицу пространственно-инвариантного импульсного отклика, описываемого формулами (9.3.2) при $J = K$. Преобразование Фурье от $\mathbf{h}_E^{(K)}$ обозначим через

$$\boldsymbol{\hbar}_E^{(K)} = (\mathbf{A}_{K^2})\,\mathbf{h}_E^{(K)}. \tag{11.2.8}$$

Полученные отсчеты преобразования расставим на главной диагонали матрицы размера $K^2 \times K^2$

$$\mathscr{H}^{(K)} = \operatorname{diag}\{\hbar_E^{(K)}(1), \ldots, \hbar_E^{(K)}(K^2)\}. \qquad (11.2.9)$$

С помощью весьма громоздких выкладок можно показать, что в спектральном пространстве матрицы оператора свертки конечных массивов и дискретизованного оператора свертки можно представить в следующем виде [4]:

$$\mathscr{D} = \mathscr{H}^{(M)}[\mathbf{P}_D \otimes \mathbf{P}_D] \qquad (11.2.10)$$

при $N = M - L + 1$ и

$$\mathscr{B} = \mathbf{P}_B \otimes \mathbf{P}_B \mathscr{H}^{(N)} \qquad (11.2.11)$$

при $N = M + L - 1$, где

$$P_D(u, v) = \frac{1}{\sqrt{M}} \frac{1 - \mathscr{W}_M^{-(u-1)(L-1)}}{1 - \mathscr{W}_M^{(u-1)} \mathscr{W}_N^{-(v-1)}}, \qquad (11.2.12а)$$

$$P_B(u, v) = \frac{1}{\sqrt{N}} \frac{1 - \mathscr{W}_N^{(v-1)(L-1)}}{1 - \mathscr{W}_M^{(u-1)} \mathscr{W}_N^{-(v-1)}}. \qquad (11.2.12б)$$

Таким образом, оба оператора свертки в спектральном пространстве содержат матрицу скалярных весовых множителей $\mathscr{H}^{(K)}$ и интерполяционную матрицу $[\mathbf{P} \otimes \mathbf{P}]$, которая позволяет связать входной вектор размера N^2 с выходным вектором размера M^2. Как правило, интерполяционная матрица $\mathbf{P} \otimes \mathbf{P}$ содержит много нулевых элементов, и поэтому свертка с использованием преобразования выполняется весьма эффективно.

Рассмотрим теперь циклическую свертку, выполняемую с переходом в спектральное пространство. С помощью рассуждений, аналогичных вышеприведенным, было показано [4], что оператор фильтрации в этом случае сводится к скалярному оператору

$$\mathscr{C} = J\mathscr{H}^{(\ )}. \qquad (11.2.13)$$

Таким образом, как видно из равенств (11.2.12) и (11.2.13), при свертке в пространстве спектров Фурье матрицу фильтра удается выразить в компактной замкнутой форме. Для других унитарных преобразований подобных выражений не найдено.

Эффективность вычисления свертки с применением преобразования Фурье определяется тем, что оператор свертки $\mathbf{C}$ имеет циклическую матрицу, а соответствующая матрица фильтра $\mathscr{C}$ является диагональной. Как можно увидеть из соотношений (11.1.6), базисные векторы преобразования Фурье фактически являются собственными векторами матрицы $\mathbf{C}$ [5]. Это не так для операции суперпозиции общего вида, а также для операции свертки с применением других унитарных преобразований. Однако во многих слу-

чаях матрицы фильтра $\mathscr{D}$, $\mathscr{B}$ и $\mathscr{E}$ содержат сравнительно мало ненулевых элементов и обработка с преобразованием позволяет уменьшить объем вычислений.

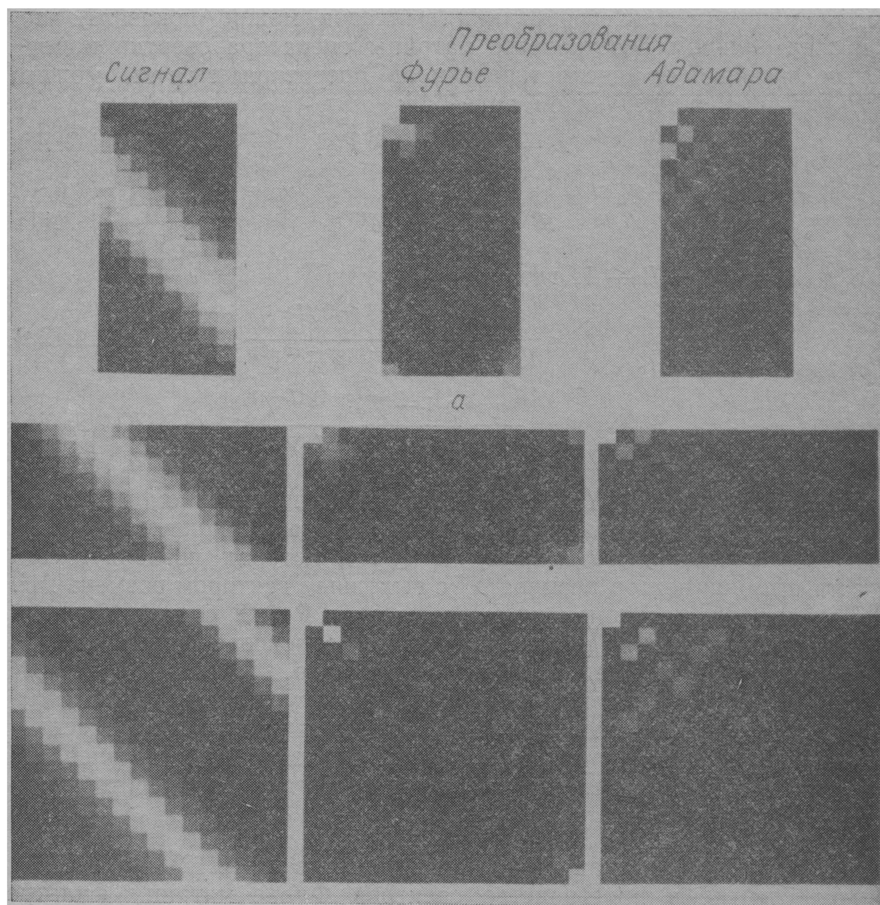


Рис. 11.2.2. Матрицы операторов свертки одномерных сигналов с использованием преобразований Фурье и Адамара.
а — конечная свертка; *б* — дискретизованная свертка; *в* — циклическая свертка.

На рис. 11.2.2 показан вид матриц спектров для трех типов оператора свертки одномерного входного вектора с гауссовым импульсным откликом с использованием преобразований Фурье и Адамара [6]. Как и ожидалось, матрицы коэффициентов преоб-

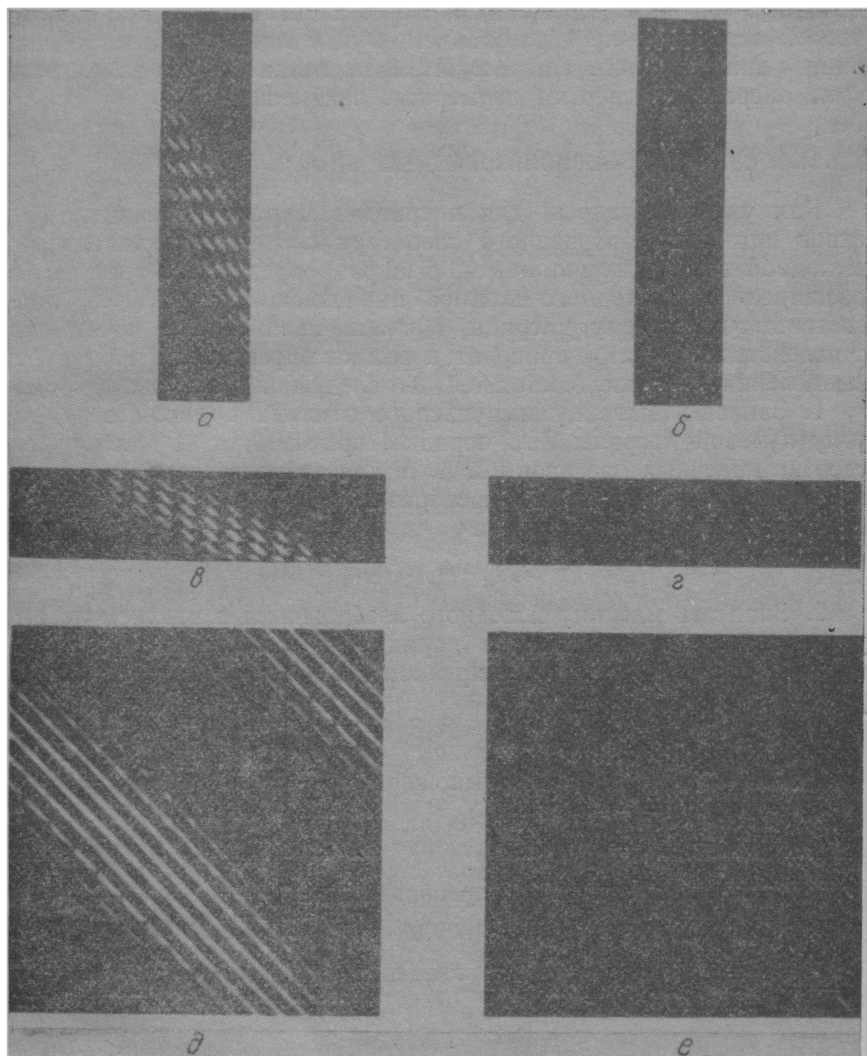


Рис. 11.2.3. Матрицы операторов свертки двумерных сигналов с использованием преобразования Фурье.

Конечная свертка: *а* — непосредственно выполняемая; *б* — с преобразованием.
Дискретизованная интегральная свертка: *в* — непосредственно выполняемая; *г* — с преобразованием.
Циклическая свертка: *д* — непосредственно выполняемая; *е* — с преобразованием.

разования оказались значительно более разреженными, чем исходные матрицы. Кроме того, легко заметить, что для циклической свертки с преобразованием Фурье матрица фильтра является диагональной. На рис. 11.2.3 показана структура матриц трех операторов свертки двумерных сигналов [4].

11.3. СВЕРТКА С ИСПОЛЬЗОВАНИЕМ БПФ

Как уже отмечалось, для оператора свертки конечных массивов или дискретизованного оператора свертки эквивалентный выходной вектор можно найти, выбирая определенные элементы расширенного выходного вектора циклической свертки $\mathbf{k}_E$ или соответствующей ему матрицы $\mathbf{K}_E$. Это положение в сочетании с равенством (11.2.13) приводит к весьма эффективной процедуре вычисления свертки, состоящей из следующих этапов:

1. Записать матрицу импульсного отклика в левом верхнем углу нулевой матрицы J-го порядка, причем $J \geqslant M$ для случая свертки конечных массивов и $J \geqslant N$ для дискретизованной свертки. Выполнить двумерное преобразование Фурье расширенной матрицы импульсного отклика

$$\mathscr{H}_E = \mathbf{A}_J \mathbf{H}_E \mathbf{A}_J. \tag{11.3.1}$$

2. Записать матрицу исходного изображения в верхнем левом углу нулевой матрицы J-го порядка и выполнить двумерное преобразование Фурье расширенной матрицы исходного изображения

$$\mathscr{F}_E = \mathbf{A}_J \mathbf{F}_E \mathbf{A}_J. \tag{11.3.2}$$

3. Выполнить скалярное умножение

$$\mathscr{K}_E(m, n) = J \mathscr{H}_E(m, n) \mathscr{F}_E(m, n), \tag{11.3.3}$$

где $1 \leqslant m \leqslant J$ и $1 \leqslant n \leqslant J$.

4. Произвести обратное преобразование Фурье

$$\mathbf{K}_E = (\mathbf{A}_J)^{-1} \mathscr{K}_E (\mathbf{A}_J)^{-1}. \tag{11.3.4}$$

5. После выбора нужных элементов сформировать искомую выходную матрицу

$$\mathbf{Q} = [\mathbf{S1}^{(M)}_J] \mathbf{K}_E [\mathbf{S1}^{(M)}_J]^T \tag{11.3.5а}$$

или

$$\mathbf{G} = [\mathbf{S2}^{(M)}_J] \mathbf{K}_E [\mathbf{S2}^{(M)}_J]^T. \tag{11.3.5б}$$

Важно, чтобы порядок J расширенных матриц $\mathscr{H}_E$ и $\mathscr{F}_E$ удовлетворял соответствующим неравенствам. При $J = N$ на левом и верхнем краях выходной матрицы будут находиться полосы ошибочных элементов, имеющие ширину в $L - 1$ отсчет (рис. 11.3.1, *а*, *б*). Они образуются в результате так называемой циклической ошибки,

связанной с неправильным применением метода БПФ для вычисления свертки. Кроме того, при выполнении свертки конечных массивов (D-свертка) будут потеряны элементы выходного массива, расположенные полосами в правой и нижней его частях. Если положить $J = M$, то при D-свертке выходная матрица будет полностью заполнена правильными отсчетами. Чтобы вычислить B-свертку

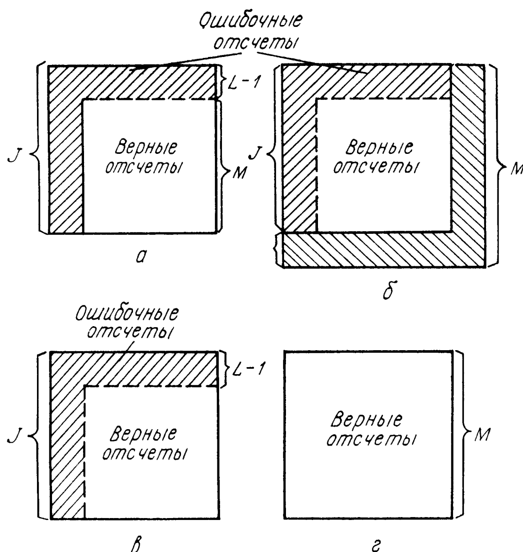


Рис. 11.3.1. Циклические ошибки. (Нижняя и правая стороны матрицы **F** отсечены.)

а — фильтрация типа B; $J = N$, $M = N - L + 1$
б — фильтрация типа D; $J = N$, $M = N + L - 1$.
в — фильтрация типа B; $J = M$, $M = N - L + 1$.
г — фильтрация типа D; $J = M$, $M = N + L - 1$.

при $J = M$, необходимо отсечь правый и нижний края исходной матрицы. Однако в результате получится, что элементы выходного массива, расположенные сверху и по левому краю, будут ошибочными.

При обработке сигналов во многих случаях оказывается, что на различные входные массивы воздействуют операторы с одним и тем же импульсным откликом и, следовательно, первый этап алгоритма (вычисление $\mathcal{H}_E$) достаточно выполнить только один раз. При использовании алгоритма БПФ для случаев прямого и обратного преобразований требуется выполнить примерно по $2J^2 \log_2 J$ арифметических операций. Скалярное умножение проводится за J^2 операций, т. е. всего требуется $J^2 (1 + 4 \log_2 J)$ операций. Если входная матрица содержит $N \times N$ элементов, выходная $M \times M$, а матрица импульсного отклика $L \times L$ элемен-

тов, то для вычисления конечной свертки требуется N^2L^2 операций, а для дискретизованной свертки M^2L^2 операций. Если размер L матрицы импульсного отклика достаточно велик по сравнению с размером N исходной матрицы, то свертка с преобразованием оказывается эффективнее прямой свертки, причем число операций может уменьшиться раз в десять или более. На рис. 11.3.2 приведен график зависимости L от N для случая,

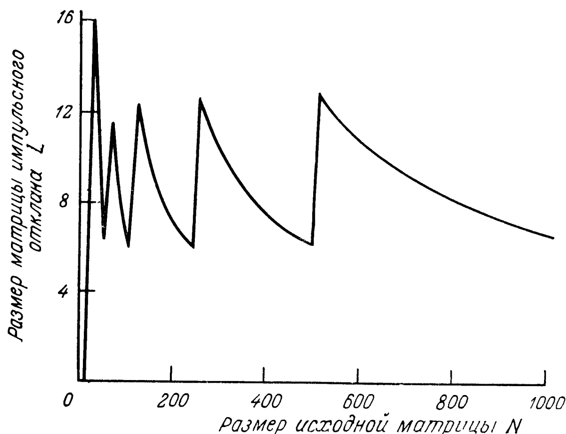


Рис. 11.3.2. Сравнение эффективности двух методов получения конечной свертки: прямого и с преобразованием Фурье.

когда оба метода вычисления свертки конечных массивов (прямой и с преобразованием Фурье) имеют одинаковую эффективность. Зубчатость графика объясняется тем, что с ростом N параметр J изменяется скачкообразно, принимая значения 64, 128, 256 и т. д.

Для вычисления свертки с преобразованием Фурье требуется меньшее число операций, чем в случае ее прямого вычисления, если длина импульсного отклика достаточно велика. Однако если обрабатываемое изображение также имеет большие размеры, то относительная эффективность метода с использованием преобразования Фурье понижается. Кроме того, при вычислении преобразования Фурье больших матриц возникают трудности, связанные с обеспечением точности расчетов. Обе проблемы удается разрешить, прибегая к блочной фильтрации изображения, когда большую матрицу подразделяют на ряд перекрывающихся блоков, обрабатываемых поочередно [2, 7—9].

На рис. 11.3.3, *а* показано, как из левого верхнего угла большой матрицы извлекается блок, содержащий $N_B \times N_B$ элементов. После свертки его с импульсным откликом, состоящим из $L \times L$ элементов, получается блок размера $M_B \times M_B$ элементов, который помещают в левый верхний угол выходной матрицы

(рис. 11.3.3, *а*). Далее из обрабатываемой матрицы извлекают следующий блок размера $N_B \times N_B$ элементоз и из него получают второй блок обработанного изображения размера $M_B \times M_B$, примыкающий к первому. Как показано на рис. 11.3.3, *б*, второй блок исходного изображения должен перекрываться с первым блоком в полосе шириной $L - 1$ элемент. Тогда обработанные блоки будут стыковаться без зазора. Этот процесс продолжается

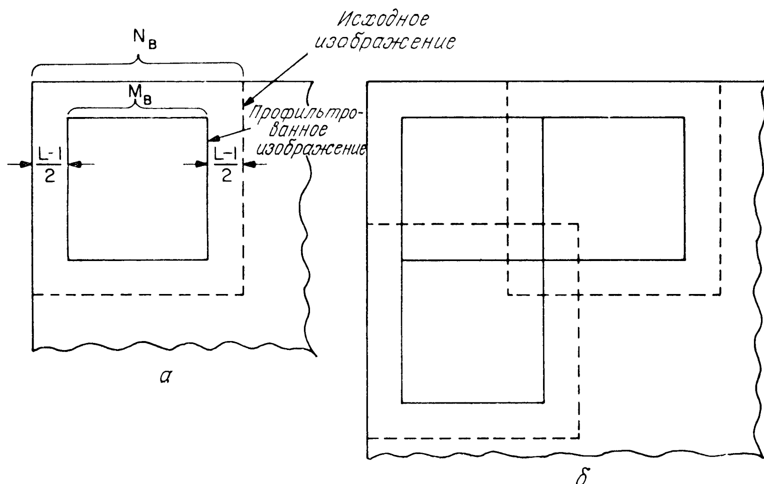


Рис. 11.3.3 Размещение блоков при блочной фильтрации изображения.
а — первый блок; *б* — вторые блоки в строке и столбце блоков.

до тех пор, пока не будут обработаны все блоки, прилегающие к верхней строке матрицы. Если в этой строке блоков последний блок окажется неполным, в него следует добавить нулевые элементы. Далее извлекают блок, находящийся в начале второй строки и перекрывающийся с блоками первой строки в полосе шириной $L - 1$ элемент. Процесс продолжается до тех пор, пока не будут определены все элементы обработанного изображения.

Для получения свертки с помощью преобразования Фурье требуется

$$O_F = N^2 + 2N^2 \log_2 N \tag{11.3.6}$$

операций. При блочной обработке, когда размер блоков равен $N_B \times N_B$, необходимо

$$O_B = R^2 [N_B^2 + 2N_B^2 \log_2 N_B] \tag{11.3.7}$$

операций, где число блоков R — округленное в бóльшую сторону значение дроби $N/[N_B + L - 1]$. Хант [9] определил, как оптимальный размер блока зависит от величины матриц исходного изображения и импульсного отклика.

11.4. ФИЛЬТРЫ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Алгоритм свертки с применением БПФ, рассмотренный в предыдущем разделе, часто используют при машинном моделировании линейных аналоговых фильтров. В данном разделе проанализированы ошибки, присущие такому методу моделирования, и описаны способы определения дискретной частотной характеристики по заданной непрерывной частотной характеристике. В целях упрощения изложения здесь описываются только одномерные сигналы.

Рассмотрим длинный одномерный непрерывный сигнал $f_C(x)$, спектр которого $\mathscr{f}_C(\omega)$ равен нулю, если $|\omega|$ больше частоты среза ω_0. Нужно найти свертку сигнала с непрерывным импульсным откликом $h_C(x)$, частотная характеристика которого $\mathscr{h}_C(\omega)$ также ограничена частотой ω_0. Как упоминалось в гл. 1, свертку можно выполнить либо в пространственной области в соответствии с соотношением

$$g_C(x) = \int_{-\infty}^{\infty} f_C(\alpha)\, h_C(x-\alpha)\, d\alpha, \tag{11.4.1a}$$

либо в частотной области

$$g_C(x) = (1/2\pi) \int_{-\infty}^{\infty} \mathscr{f}_C(\omega)\, \mathscr{h}_C(\omega) \exp\{i\omega x\}\, d\omega. \tag{11.4.1б}$$

В гл. 9 описана методика дискретизации интеграла свертки (11.4.1). Непрерывный импульсный отклик $h_C(x)$ необходимо подвергнуть усечению, умножив его на весовую функцию $y(x)$. В результате получается взвешенный импульсный отклик

$$b_C(x) = h_C(x)\, y(x), \tag{11.4.2}$$

где $y(x) = 0$ при $|x| > T$. Весовая функция уменьшает эффекты, связанные с усечением. Интеграл свертки аппроксимируется выражением.

$$g_C(x) = \int_{x-T}^{x+T} f_C(\alpha)\, b_C(x-\alpha)\, d\alpha. \tag{11.4.3}$$

Затем в $2J+1$ точке берут отсчеты выходного сигнала с интервалом $\Delta = \pi/\omega_0$, а непрерывный интеграл заменяют суммой с тем же шагом Δ. В результате получается дискретное представление

$$g_C(j\Delta) = \sum_{k=j-K}^{j+K} f_C(k\Delta)\, b_C[(j-k)\Delta], \tag{11.4.4}$$

где K — целое число, ближайшее к значению дроби T/Δ.

Для вычисления суммы (11.4.4) с помощью дискретного преобразования Фурье можно следующим образом применить алгоритм, описанный в разд. 11.3. На первом этапе в качестве первых

$L = 2K + 1$ элементов J-элементной последовательности берутся отсчеты взвешенного импульсного отклика, а в качестве последующих $J - L$ элементов — нули. Таким образом,

$$\underbrace{b_D(\ell) = b_C(-K), \ldots, b_C(0), \ldots, b_C(K),}_{L \text{ элементов}} 0, \ldots, 0, \qquad (11.4.5)$$

где $0 \leqslant \ell \leqslant P - 1$. Элементы последовательности $b_D(\ell)$ можно получить из непрерывного импульсного отклика $h_C(x)$ и весовой функции $y(x)$, дискретизируя произведение этих функций, т. е.

$$b_D(\ell) = y(x) h_C(x) \delta(x + K\Delta - \ell\Delta). \qquad (11.4.6)$$

На следующем этапе вычисляется дискретное преобразование Фурье $b_D(\ell)$ в P точках

$$\mathscr{b}_D(u) = \frac{1}{\sqrt{P}} \sum_{l=0}^{P-1} b_D(\ell) \exp\left\{-\frac{2\pi i}{P} \ell u\right\}, \qquad (11.4.7)$$

где $0 \leqslant u \leqslant P - 1$. После подстановки выражения для $b_D(\ell)$ в формулу (11.4.7) и преобразований получается, что дискретная частотная характеристика фильтра связана с аналоговой частотной характеристикой $\mathscr{h}_C(\omega)$ и спектром Фурье $\mathscr{y}(\omega)$ весовой функции соотношением

$$\mathscr{b}_D(u) = \frac{1}{4\pi^2} \frac{1}{\sqrt{P}} \exp\left\{-i\pi\left(\frac{L-1}{P}\right) u\right\} \mathscr{b}_C\left(\frac{2\omega_0}{P} u\right), \qquad (11.4.8а)$$

$$\mathscr{b}_D(P - u) = \mathscr{b}_D^*(u), \qquad (11.4.8б)$$

где $u = 0, 1, \ldots, P/2$, а $\mathscr{b}_C(\omega) = \mathscr{h}_C(\omega) * \mathscr{y}(\omega)$.

Равенства (11.4.8) задают искомую связь между дискретной и непрерывной частотными характеристиками. Если непрерывная частотная характеристика $\mathscr{h}_C(\omega)$ и спектр Фурье $\mathscr{y}(\omega)$ весовой функции известны и заданы в аналитическом виде, то в принципе отсчеты дискретной частотной характеристики можно получить, выполнив аналитическим путем свертку (11.4.8б) и находя числовые значения полученной при этом функции в точках $(2\pi u/P\Delta)$ для каждого значения параметра u. На практике часто не удается вычислить свертку аналитически, особенно для двумерных сигналов, и оказывается проще выполнить обратное преобразование Фурье частотной характеристики $\mathscr{h}_C(\omega)$ для получения аналитического выражения импульсного отклика и затем взять отсчеты $b_D(\ell)$ в соответствии с формулой (11.4.6). Можно использовать и другой подход: согласно равенству (11.4.8), взять отсчеты дискретного обратного преобразования Фурье $\mathscr{h}_C(2\pi u/P\Delta)$, умножить расширенную последовательность отсчетов импульсного отклика на весовую функцию, а затем, выполнив дискретное преобразование Фурье, получить $\mathscr{b}_D(u)$.

Умножение на весовую функцию, которое можно выполнить согласно равенству (11.4.6) или неявно в спектральной области с помощью соотношения (11.4.8), совершенно необходимо, если нужно подавить описанные в гл. 9 циклические ошибки. При фильтрации изображений обычно делается типичная ошибка, когда в качестве дискретного импульсного отклика берется последовательность отсчетов непрерывного импульсного отклика. Тогда в общем случае все J элементов соответствующего расширенного дискретного импульсного отклика будут ненулевыми, т. е. длина L дискретного импульсного отклика, «погружаемого» в расширенный вектор (11.4.5), неявно будет принята равной J. Поэтому все профильтрованные отсчеты будут искажены из-за циклической ошибки.

Известно [10—12] несколько видов весовых функций, пригодных для использования при дискретной линейной фильтрации. Некоторые из них, чаще всего встречающиеся на практике, приведены в табл. 11.4.1, а их графики — на рис. 11.4.1. На рис. 11.4.2 представлены соответствующие спектры, состоящие из главного лепестка и набора боковых лепестков, амплитуды которых обычно уменьшаются с ростом частоты. Анализ выражения

Таблица 11.4.1

Весовые функции

Функция	Определение
Прямоугольная	$w(n) = 1, \quad 0 \leqslant n \leqslant L-1$
Бартлетта (треугольная)	$w(n) = \begin{cases} \dfrac{2n}{L-1}, & 0 \leqslant n < \dfrac{L-1}{2}, \\ 2 - \dfrac{2n}{L-1}, & \dfrac{L-1}{2} \leqslant n \leqslant L-1 \end{cases}$
Ханна	$w(n) = \dfrac{1}{2}\left[1 - \cos\left(\dfrac{2\pi n}{L-1}\right)\right], \quad 0 \leqslant n \leqslant L-1$
Хемминга	$w(n) = 0{,}54 - 0{,}46\cos\left(\dfrac{2\pi n}{L-1}\right), \quad 0 \leqslant n \leqslant L-1$
Блэкмана	$w(n) = 0{,}42 - 0{,}5\cos\left(\dfrac{2\pi n}{L-1}\right) + 0{,}08\cos\left(\dfrac{4\pi n}{L-1}\right),$ $0 \leqslant n \leqslant L-1$
Кайзера	$\dfrac{I_0\{\omega_a[((L-1)/2)^2 - [n - ((L-1)/2)]^2]^{1/2}\}}{I_0\{\omega_a((L-1)/2)\}},$ $0 \leqslant n \leqslant L-1,$ $I_0\{\cdot\}$ — модифицированная функция Бесселя нулевого порядка первого рода

(11.4.8) показывает, каким образом форма весовой функции и ее спектра влияет на выходной сигнал. Форма главного лепестка спектра весовой функции определяет искажение спектра сигнала в диапазоне от 0 до ω_0. Боковые лепестки вызывают наложение спектров, поскольку взвешенный импульсный отклик $\mathscr{b}_C(\omega)$ имеет спектр неограниченной ширины. Плавные весовые функции

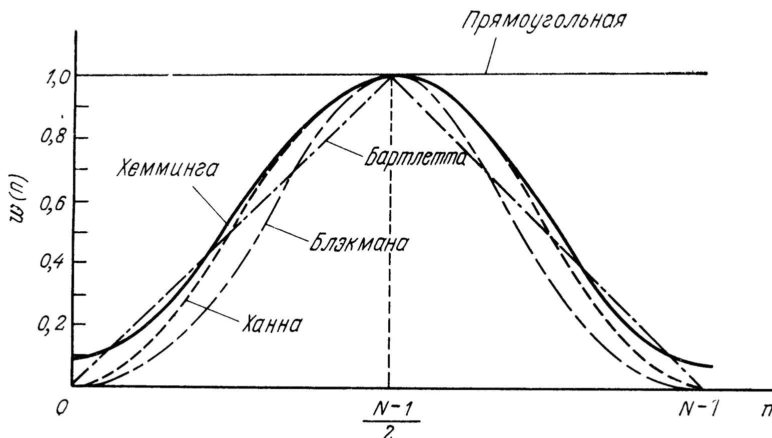


Рис. 11.4.1. Одномерные весовые функции [10].

имеют слабые боковые лепестки, что уменьшает ошибки наложения спектров, однако главный лепесток в этих случаях оказывается более широким, что приводит к сглаживанию спектра сигнала. При проектировании фильтров приходится находить компромисс между этими двумя видами ошибок. Обе ошибки можно уменьшить, увеличивая длину взвешенного импульсного отклика, однако при этом либо сокращается длина обработанного сигнала, либо увеличивается объем вычислений.

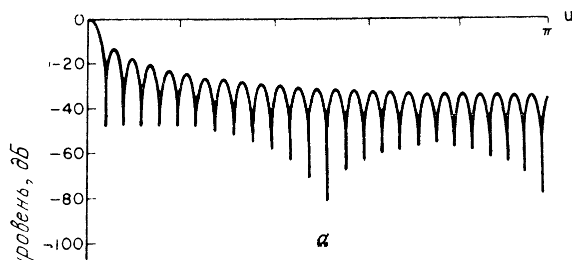
11.5. ПСЕВДООБРАЩЕНИЕ С ПРЕОБРАЗОВАНИЕМ

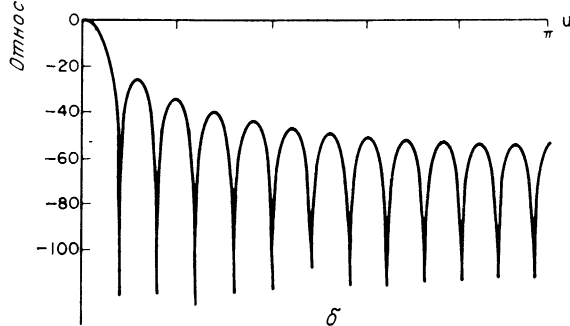
Обработку с использованием преобразований можно с успехом применить для получения обобщенных обратных матриц. Как видно из соотношения (11.1.6б), матрица $\mathscr{T}$ линейного оператора в спектральном пространстве связана с произвольной матрицей **T** размера $M^2 \times N^2$ следующим образом.

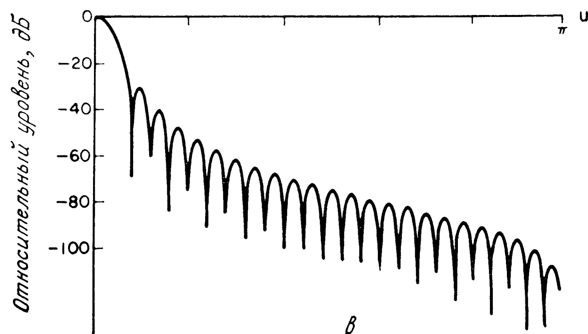
$$\mathscr{T} = (\mathbf{A}_{M^2})\,\mathbf{T}\,(\mathbf{A}_{N^2})^{-1}. \qquad (11.5.1)$$

В тех же обозначениях обобщенная обратная матрица $\mathscr{T}^-$ определяется как

$$\mathscr{T}^- = (\mathbf{A}_{N^2})\,\mathbf{T}^-\,(\mathbf{A}_{M^2})^{-1}, \qquad (11.5.2)$$


0
-20
-40
-60
-80
-100
u
π
a


Относительный уровень, дБ
0
-20
-40
-60
-80
-100
u
π
б


Относительный уровень, дБ
0
-20
-40
-60
-80
-100
u
π
в

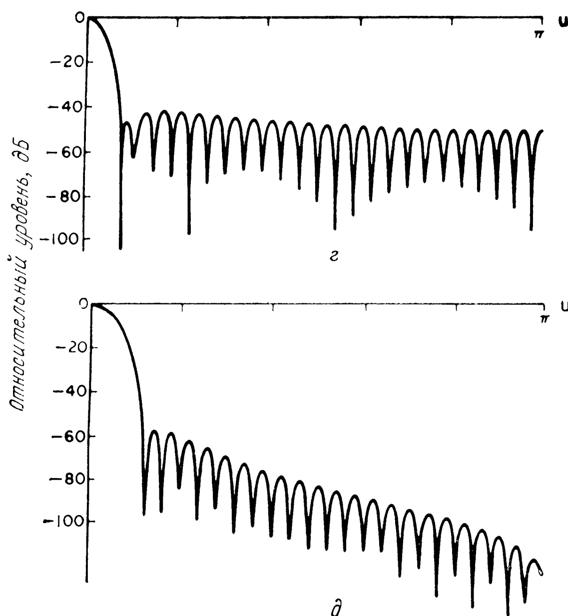


Рис. 11.4.2. Спектры одномерных весовых функций [10].

а — прямоугольной; *б* — треугольной (Бартлетта), *в* — Ханна; *г* — Хемминга; *д* — Блэкмана.

где $\mathbf{T}^-$ — матрица размера $N^2 \times M^2$. Соотношения (11.5.1) и (11.5.2) являются самосогласованными, поскольку известно, что для произвольной матрицы $\mathbf{Z}$ и унитарных матриц $\mathbf{P}$ и $\mathbf{R}$ выполняется равенство $(\mathbf{PZR})^- = \mathbf{R}^{*T}\mathbf{Z}^-\mathbf{P}^{*T}$ [13, стр. 100].

Если ранг матрицы $\mathbf{T}$ равен N^2, то, согласно формуле (8.3.5),

$$\mathbf{T}^- = (\mathbf{T}^T\mathbf{T})^{-1}\,\mathbf{T}^T \tag{11.5.3}$$

и нетрудно показать, что

$$\mathcal{T}^- = (\mathcal{T}^{*T}\mathcal{T})^{-1}\,\mathcal{T}^{*T}. \tag{11.5.4}$$

В противоположном случае, когда ранг матрицы $\mathbf{T}$ равен M^2, из формулы (8.3.6) следует, что

$$\mathbf{T}^- = \mathbf{T}^T\,(\mathbf{T}\mathbf{T}^T)^{-1} \tag{11.5.5}$$

и обобщенная обратная матрица преобразования в спектральном пространстве удовлетворяет соотношению

$$\mathcal{T}^- = \mathcal{T}^{*T}\,(\mathcal{T}\mathcal{T}^{*T})^{-1}. \tag{11.5.6}$$

Укажем для примера, что обобщенные обратные матрицы относительно матриц максимального ранга, соответствующих опера-

торам суперпозиции $\mathscr{D}$, $\mathscr{B}$ и $\mathscr{C}$, определяются равенствами

$$\mathscr{D}^{-} = (\mathscr{D}^{*T}\mathscr{D})^{-1}\,\mathscr{D}^{*T}, \tag{11.5.7а}$$

$$\mathscr{B}^{-} = \mathscr{B}^{*T}\,(\mathscr{B}\mathscr{B}^{*T})^{-1}, \tag{11.5.7б}$$

$$\mathscr{C}^{-} = \mathscr{C}^{-1}. \tag{11.5.7в}$$

На рис. 11.5.1 представлены распечатки обобщенных обратных матриц для операторов свертки одномерных сигналов с исполь-

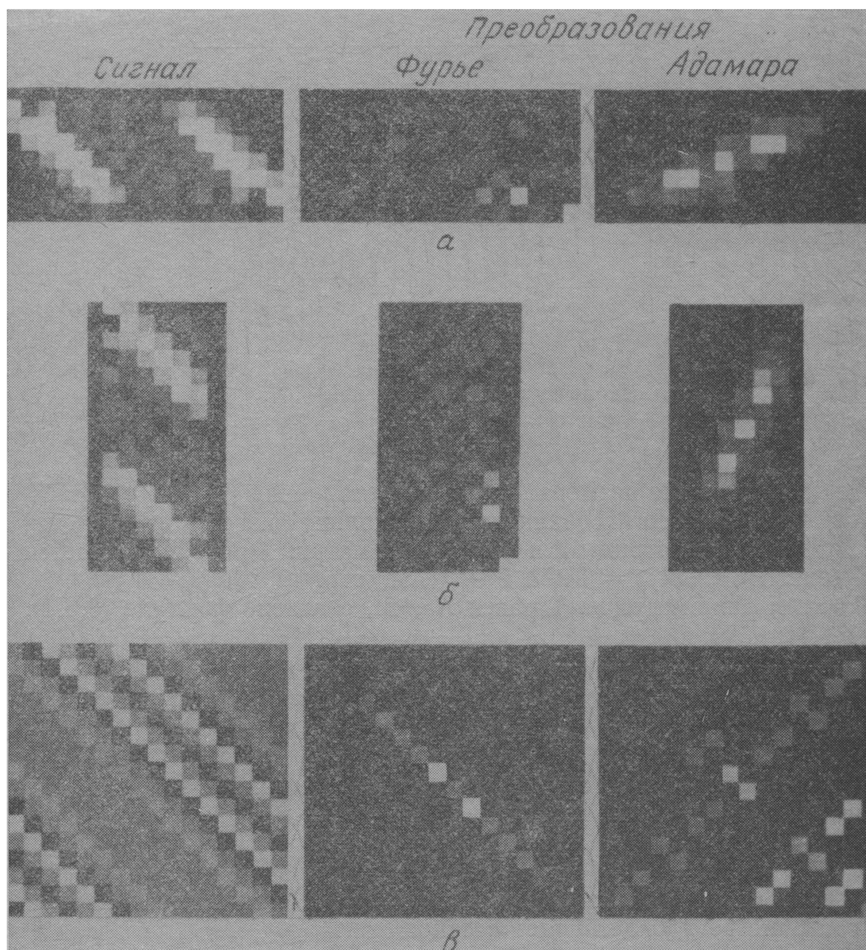


Рис. 11.5.1. Обобщенные обратные матрицы операторов свертки одномерных сигналов с использованием преобразований Фурье и Адамара.

а — конечная свертка; *б* — дискретизованная интегральная свертка *в* — циклическая свертка.

зованием преобразований Фурье и Адамара. Хорошо заметно, что эти матрицы более разрежены, чем исходные матрицы. Более того, обобщенная обратная матрица оператора циклической свертки с преобразованием Фурье является диагональной, поскольку, как следует из равенства (11.2.13), матрица $\mathscr{C}$ — диагональная.

11.6. РЕКУРСИВНАЯ ФИЛЬТРАЦИЯ

В предыдущих разделах данной главы обработка с преобразованием рассматривалась как косвенный метод выполнения двумерной линейной обработки. Было показано, что для обработки с преобразованием часто требуется гораздо меньше арифметических операций, чем при использовании стандартных методов. В данном разделе будет рассмотрен другой способ линейной обработки, называемый рекурсивной фильтрацией [14—16]. Иногда рекурсивная фильтрация оказывается даже более эффективной, чем обработка с преобразованием. Кроме того, в этом случае для хранения данных требуется ЗУ меньшей емкости, чем при обработке с преобразованием.

Рекурсивная фильтрация основывается на рекуррентном соотношении между входными и выходными переменными системы. Для одномерных сигналов подобное рекуррентное соотношение имеет следующий вид [14]:

$$r(m) = \sum_{j=1}^{J} a(j) f(m-j+1) - \sum_{k=2}^{K}{}' \, b(k) r(m-k+1), \quad (11.6.1)$$

где $f(n)$, $n = 1, 2, \ldots, N$ — отсчеты входной последовательности, $r(m)$, $m = 1, 2, \ldots, M$ — отсчеты выходной последовательности, а $b(k)$ и $a(j)$ — весовые множители. Ключевой момент здесь в том, что m-й элемент выходной последовательности зависит не только от последнего и $J-1$ предпоследних элементов входной последовательности, но и от $K-1$ предыдущих элементов выходной последовательности.

Большинство методов синтеза и анализа рекурсивных фильтров основано на применении Z-преобразования. По определению [17, 18] Z-преобразование N-элементной последовательности $f(n)$ дает образ

$$\mathscr{f}(z) = \sum_{n=1}^{N} f(n) z^{1-n}. \quad (11.6.2)$$

Нетрудно показать, что Z-преобразование выражения (11.6.1) дает образ

$$\mathscr{r}(z) = \frac{a(z)}{\mathscr{b}(z)} \mathscr{f}(z), \quad (11.6.3)$$

где $a(z)$, $b(z)$ и $f(z)$ — образы соответствующих последовательностей.

Двумерная рекурсивная фильтрация основана на следующем рекуррентном соотношении между входным и выходным массивами [19, 20]:

$$R(m_1, m_2) = \sum_{j_1=1}^{1} \sum_{j_2=1}^{J_2} A(j_1, j_2) F(m_1 - j_1 + 1, m_2 - j_2 + 1) -$$

$$- \sum_{k_1=2}^{K_1} \sum_{k_2=2}^{K_2} B(k_1, k_2) R(m_1 - k_1 + 1, m_2 - k_2 + 1), \quad (11.6.4)$$

где $F(n_1, n_2)$ — входной массив из $N_1 \times N_2$ элементов, $R(m_1, m_2)$ — выходной массив из $M_1 \times M_2$ элементов, а $B(k_1, k_2)$ и $A(j_1, j_2)$ — весовые множители. Предполагается, что процесс рекурсивной фильтрации начинается с левого верхнего угла входного массива. С помощью двумерного Z-преобразования из равенства (11.6.4) получается

$$\mathcal{R}(z_1, z_2) = \frac{\mathcal{A}(z_1, z_2)}{\mathcal{B}(z_1, z_2)} \mathcal{F}(z_1, z_2), \quad (11.6.5)$$

где $\mathcal{A}(z_1, z_2)$, $\mathcal{B}(z_1, z_2)$ и $\mathcal{F}(z_1, z_2)$ — двумерные образы соответствующих массивов. Так, например,

$$\mathcal{A}(z_1, z_2) = \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} A(n_1, n_2) z_1^{(1-n_1)} z_2^{(1-n_2)}. \quad (11.6.6)$$

При синтезе рекурсивных фильтров требуется выбрать такие массивы весовых множителей A и B, чтобы выходной массив $R(m_1, m_2)$ оказался эквивалентным массиву, получаемому в результате свертки функции $F(n_1, n_2)$, описывающей исходное изображение, с заданным импульсным откликом $H(\ell_1, \ell_2)$. Как правило, точного совпадения массивов добиться не удается и при синтезе фильтров приходится пользоваться приближенными методами [21, 22]. При этом возникает вопрос об устойчивости рассчитанного рекурсивного фильтра. Если фильтр неустойчив, то ошибки округления или шум, присутствующий во входном массиве, могут в ходе обработки не ослабляться, а увеличиваться до очень большого уровня. Рекурсивный фильтр устойчив [20], если коэффициенты $K(j_1, j_2)$ разложения частотной характеристики фильтра в ряд по степеням переменных z_1 и z_2

$$\frac{\mathcal{A}(z_1, z_2)}{\mathcal{B}(z_1, z_2)} \equiv \mathcal{K}(z_1, z_2) = \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} K(j_1, j_2) z_1^{1-j_1} z_2^{1-j_2} \quad (11.6.7)$$

являются абсолютно суммируемыми, т. е. удовлетворяют условию

$$\sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} |K(j_1, j_2)| < \infty. \quad (11.6.8)$$

Разработано несколько методов проверки рекурсивных фильтров на устойчивость [23—25]. При обработке изображения согласно равенству (11.6.4) требуется выполнить

$$N_R = (J_1J_2 + K_1K_2)\, N_1N_2 \qquad (11.6.9)$$

арифметических операций. Здесь J_i и K_i — размеры массивов весовых множителей для входного и выходного изображений, а N_i — размеры входного изображения. Если изображения и массивы весовых множителей — квадратные (т. е. $J_1 = J_2 = J$, $K_1 = K_2 = K$ и $N_1 = N_2 = N$), то при рекурсивной фильтрации нужно выполнить

$$N_R = (J^2 + K^2)\, N^2 \qquad (11.6.10)$$

операций. Для сравнения укажем, что для получения конечной свертки (см. разд. 9.3) требуется

$$N_C = L^2N^2 \qquad (11.6.11)$$

операций, где L — размер импульсного отклика. Как показано в разд. 11.2, для получения свертки с применением быстрого преобразования Фурье необходимо примерно

$$N_F = N^2 (1 + 4 \log_2 N) \qquad (11.6.12)$$

операций. Степень относительной эффективности трех типов обработки зависит от размеров массива импульсного отклика [27]. Если эти размеры невелики, то для прямого метода получения свертки и рекурсивной фильтрации требуется почти одинаковое число арифметических операций. В этом случае сравнить эффективность обоих методов с эффективностью получения свертки с использованием БПФ можно с помощью графика на рис. 11.3.2. При больших размерах массива импульсного отклика быстрый метод получения свертки оказался гораздо эффективнее прямого метода. Сравнивая число операций в равенствах (11.6.10) и (11.6.12), можно показать, что при больших размерах импульсного отклика рекурсивная фильтрация оказывается эффективнее быстрого метода получения свертки, если площади массивов коэффициентов рекурсивного фильтра удовлетворяют неравенству

$$J^2 + K^2 \leq c \log_2 N, \qquad (11.6.13)$$

где c — постоянная, принимающая значения от 1 до 20, причем ее величина зависит от вида использованного алгоритма БПФ и от того, насколько вычисления с комплексными числами происходят медленнее, чем с действительными.

ЛИТЕРАТУРА

1. Pratt W. K , Generalized Wiener Filtering Computation Techniques, *IEEE Trans. Computers*, **C-21**, 7, 636—641 (July 1972).
2. Stockham T. G , Jr., High Speed Convolution and Correlation, Proceedings Spring Joint Computer Conference, 1966, pp. 229—233.
3. Gentleman W. M., Sande G., Fast Fourier Transforms — for Fun and Profit, Proceedings Fall Joint Computer Conference, 1966, pp. 563—578.
4. Pratt W. K., Vector Formulation of Two-Dimensional Signal Processing Operations, *J. Computer Graphics and Image Processing* (March 1975) (Academic Press, New York).
5. Hunt B. R., A Matrix Theory Proof of the Discrete Convolution Theorem, *IEEE Trans. Audio and Electroacoustics*, **AU-19**, 4, 285—288 (December 1973).
6. Pratt W. K., Transform Domain Signal Processing Techniques, Proceedings National Electronics Conference, Chicago, Illinois, 1974.
7. Helms H. D., Fast Fourier Transform Method of Computing Difference Equations and Simulating Filters, *IEEE Trans. Audio and Electroacoustics*, **AU-15**, 2, 85—90 (June 1967).
8. Ekstrom M. P., Algazi V. R., Optimum Design of Two-Dimensional Nonrecursive Digital Filters, Proceedings 4th Asilomar Conference on Circuits and Systems, Pacific Grove, California, November 1970.
9. Hunt B. R., Computational Considerations in Digital Image Enhancement, Proceedings Conference on Two-Dimensional Signal Processing, University of Missouri, Columbia, Missouri, October 1971.
10. Oppenheim A. V., Schafer R. W., Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N. J., 1975. [Имеется перевод: Оппенгейм А. В., Шафер Р. В., Цифровая обработка сигналов. — М.: Связь, 1979.]
11. Blackman R. B., Tukey J. W., The Measurement of Power Spectra, Dover, New York, 1958.
12. Kaiser J. F., Digital Filters, Chapter 7 in Systems Analysis by Digital Computer, Kuo F. F., Kaiser J. F., Eds., Wiley, New York, 1966. [Имеется перевод: Кайзер Дж., Цифровые фильтры. Приложение к книге: Голд Б., Рэйдер Ч., Цифровая обработка сигналов. — М.: Советское радио, 1973.]
13. Graybill F. A., Introduction to Matrices with Applications in Statistics, Wadsworth Publishing, Belmont, Calif., 1969.
14. Childers D. G., Durling A., Digital Filtering and Signal Processing, West Publishing, St. Paul, Minn., 1975.
15. Oppenheim A. V., Schafer R. W., Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N. J., 1975.
16. Rabiner L. R., Gold B., Theory and Application of Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N. J., 1975. [Имеется перевод: Рабинер Л., Гоулд Б., Теория и применение цифровой обработки сигналов. — М.: Мир, 1978.]
17. Ragazzini J. R., Franklin G. F., Sampled Data Control Systems, McGraw-Hill, New York, 1958.
18. Jury E. I., Theory and Applications of the Z-Transform Method, Wiley, New York, 1964.
19. Shanks J. L., Two-Dimensional Recursive Filters, SWIEECO Record, Vol. 21, 19E1, 1969.
20. Farmer C. H., Gooden D. S., Rotation and Stability of a Recursive Digital Filter, Proceedings Two-Dimensional Signal Processing Conference, University of Missouri, Columbia, Missouri, October 1971.
21. Hu J. V., Rabiner L. R., Design Techniques for Two-Dimensional Digital Filters, *IEEE Trans. Audio and Electroacoustics*, **AU-20**, 4, 249—257 (October 1972).
22. Dudgeon D. E., Two-Dimensional Recursive Filter Design Using Differential Correction, *IEEE Trans. Acoustics, Speech, and Signal Processing*, **ASSP-23**, 3, 264—267 (June 1975).

23. Shanks J. L., Treitel S., Justice J. H., Stability and Synthesis of Two-Dimensional Recuisive Filters, *IEEE Trans. Audio and Electroacoustics*, **AU-20**, 2, 115—128 (June 1972).
24. Huang T. S , Stability of Two-Dimensional Recursive Filters, *IEEE Trans. Audio and Electroacoustics*, **AU-20**, 2, 158—163 (June 1972).
25. Anderson B. D., Jury E. I., Stability Test for Two-Dimensional Recursive Filters, *IEEE Trans. Audio and Electroacoustics*, **AU-21**, 4, 366—372 (August 1973).
26. Pistor P., Stability Criterion for Recursive Filters, *IBM J. Res. Devel.*, **18**, 1, 59—71 (January 1974).
27. Hall E. L., A Comparison of Computations for Spatial Frequency Filtering, *Proc. IEEE*, **60**, 7, 887—891 (July 1972). [Имеется перевод: Холл. Сравнение трех методов цифровой пространственной фильтрации. — ТИИЭР, 1972, т. 60, № 7.]

ОГЛАВЛЕНИЕ

Уильям Прэтт

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

Старший научный редактор Л. П. Якименко
Младший научный редактор Е. П. Орлова
Художник В. И. Харламов
Художественный редактор Л. Е. Безрученков
Технический редактор Н. Д. Толстякова
Корректор Е. Г. Литвак

ИБ № 2371

Сдано в набор 25.03.81.
Подписано к печати 29.10.81.
Формат 60×90 1/16.
Бумага типографская № 2.
Гарнитура литературная. Печать высокая.
Объем 9,75 бум. л. Усл. печ. л. 19,50.
Усл. кр.-отт. 19,50.
Уч.-изд. л. 16,61. Изд. № 20/0958.
Тираж 10 000 экз. Заказ № 596. Цена 1 р. 50 к.

ИЗДАТЕЛЬСТВО «МИР»
129820, Москва, И-110, ГСП,
1-й Рижский пер., 2.

Ленинградская типография № 6 ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли
193144, Ленинград, ул. Моисеенко, 10

1р. 50к.

У. Прэтт

Цифровая обработка изображений

1